

ENRICO ONOFRI

LEZIONI SULLA

TEORIA DEGLI OPERATORI LINEARI

1	2	3	4	5	0	0	0	0	0	0	0	0	0	0	0	0	5	4	3	2	1	
2	4	6	8	4	5	0	0	0	0	0	0	0	0	0	0	0	5	4	8	6	4	2
3	6	9	6	8	4	5	0	0	0	0	0	0	0	0	0	5	4	8	6	9	6	3
4	8	6	9	6	8	4	5	0	0	0	0	0	0	0	5	4	8	6	9	6	8	4
5	4	8	6	9	6	8	4	5	0	0	0	0	0	5	4	8	6	9	6	8	4	5
0	5	4	8	6	9	6	8	4	5	0	0	0	5	4	8	6	9	6	8	4	5	0
0	0	5	4	8	6	9	6	8	4	5	0	5	4	8	6	9	6	8	4	5	0	0
0	0	0	5	4	8	6	9	6	8	4	10	4	8	6	9	6	8	4	5	0	0	0
0	0	0	0	5	4	8	6	9	6	13	8	13	6	9	6	8	4	5	0	0	0	0
0	0	0	0	0	5	4	8	6	14	10	16	10	14	6	8	4	5	0	0	0	0	0
0	0	0	0	0	0	5	4	13	10	17	12	17	10	13	4	5	0	0	0	0	0	0
0	0	0	0	0	0	0	10	8	16	12	18	12	16	8	10	0	0	0	0	0	0	0
0	0	0	0	0	0	5	4	13	10	17	12	17	10	13	4	5	0	0	0	0	0	0
0	0	0	0	0	5	4	8	6	14	10	16	10	14	6	8	4	5	0	0	0	0	0
0	0	0	0	5	4	8	6	9	6	13	8	13	6	9	6	8	4	5	0	0	0	0
0	0	0	5	4	8	6	9	6	8	4	10	4	8	6	9	6	8	4	5	0	0	0
0	0	5	4	8	6	9	6	8	4	5	0	5	4	8	6	9	6	8	4	5	0	0
0	5	4	8	6	9	6	8	4	5	0	0	0	5	4	8	6	9	6	8	4	5	0
5	4	8	6	9	6	8	4	5	0	0	0	0	0	5	4	8	6	9	6	8	4	5
4	8	6	9	6	8	4	5	0	0	0	0	0	0	0	5	4	8	6	9	6	8	4
3	6	9	6	8	4	5	0	0	0	0	0	0	0	0	0	5	4	8	6	9	6	3
2	4	6	8	4	5	0	0	0	0	0	0	0	0	0	0	0	5	4	8	6	4	2
1	2	3	4	5	0	0	0	0	0	0	0	0	0	0	0	0	0	5	4	3	2	1

UNIVERSITÀ DEGLI STUDI DI PARMA

Seconda Edizione, 2009

Ad Alessio, Bartolomeo, Chiara, Cristian e Liliana.

Author:

Enrico Onofri (1946-.)


Dipartimento di Fisica


Università di Parma

43121 Parma, Italy

Il mio ringraziamento a **Donald E. Knuth** (\TeX), **Leslie Lamport** (\LaTeX), **Richard Stallman** (gnu), **David Reitter** (aquamacs-emacs) e al mondo *Mac* e *Open Source* che hanno reso possibile quest'opera. Per i problemi che hanno richiesto calcolo numerico ci siamo avvalsi del linguaggio **matlab** (THE MATHWORKSTM) e per calcolo simbolico di **Mathematica** (WOLFRAM RESEARCH). L'archivio di codici **matlab** e **Mathematica** e le versioni aggiornate del testo sono reperibili sul sito

<http://www.fis.unipr.it/home/enrico.onofri>

 *In copertina: una matrice “iperbolica”: per saperne di più risolvere l'esercizio 46 (parte II).*

Composto su  utilizzando **TeXLive-2008** nel formato **amsbook**, con packages aggiuntivi **bbm**, **esint**, **fancybox**, **fancyhdr**, **fancyvrb**, **makeidx**, **manfnt**, **marginnote**, **mathtools**, **MnSymbol** e **xcolor** .

© *Università di Parma, 2009*

Copying and printing this document is granted to anyone for scholar purposes, provided its content is not modified in any way.

Indice

<i>Prefazione</i>	vii
<i>Notazioni</i>	viii
<i>Prefazione alla prima edizione</i>	ix
I - Introduzione	1
Capitolo 1. Spazi lineari in Meccanica Classica	3
1.1. Il tensore d'inerzia	3
1.2. Modi normali di due oscillatori accoppiati	6
1.3. Modi normali di n oscillatori accoppiati	8
1.4. Spazi a infinite dimensioni	11
1.4.1. Un sistema con n particelle e $n \rightarrow \infty$	11
1.4.2. Un'equazione integrale	14
Problemi	17
Soluzioni	19
II - Operatori Lineari in spazi a dimensione finita	21
Capitolo 2. Spazi Lineari	23
2.1. Spazi lineari reali	23
2.1.1. Assiomi	23
2.1.2. Basi in uno spazio lineare	24
2.2. Spazi lineari complessi	27
2.3. Spazi lineari metrici	27
2.3.1. Cambiamenti di base e matrici unitarie	29
2.4. Funzionali lineari e spazio duale	31
2.4.1. La notazione di Dirac	32
Capitolo 3. Operatori Lineari	35
3.1. Definizioni fondamentali	35
3.1.1. Cambiamenti di base	36
3.1.2. Algebra degli operatori lineari	37
3.1.3. Funzioni di operatori lineari	37
3.1.4. L'operatore inverso	41
3.2. Teoria spettrale	42
3.2.1. Proprietà dello spettro	44
3.2.2. La decomposizione spettrale	45
3.3. Operatori autoaggiunti, unitari, normali	52

3.3.1. Operatori unitari	53
3.3.2. Operatori normali	55
3.4. Funzioni di più operatori lineari	55
3.4.1. Insiemi completi di operatori autoaggiunti	55
3.4.2. Insiemi irriducibili di operatori	56
3.4.3. Funzioni di operatori non-commutanti	57
3.4.4. Teoria delle perturbazioni	59
3.4.5. Soluzione di sistemi lineari	66
3.4.6. Rappresentazione di interazione	67
Appendice II	69
3.5. Polinomi Ortogonali	69
3.5.1. Definizioni	69
3.5.2. Relazioni di ricorrenza	71
3.5.3. Zeri dei polinomi ortogonali	72
3.5.4. Formule di quadratura di Gauss	73
3.5.5. Polinomi classici	75
3.6. Disuguaglianze fondamentali	76
Problemi	81
Soluzioni	95
III - Operatori lineari nello spazio di Hilbert	105
Capitolo 4. Spazio di Hilbert	107
4.1. Spazi lineari ad infinite dimensioni	107
4.1.1. Definizioni fondamentali	107
4.1.2. Funzionali lineari	112
4.2. Operatori lineari nello spazio di Hilbert	114
4.2.1. Operatori lineari chiusi, continui, limitati	114
4.2.2. Aggiunto Hermitiano di operatori limitati	117
4.2.3. Operatori isometrici e unitari	117
4.2.4. Aggiunto Hermitiano per operatori non limitati	118
4.2.5. Operatori differenziali in $L_2(a, b)$	120
4.3. Teoria spettrale	124
4.3.1. Prime proprietà ed esempi	125
4.3.2. Proprietà analitiche del risolvente	127
4.3.3. Gli operatori canonici della Meccanica Quantistica	128
4.3.4. La teoria spettrale per operatori autoaggiunti	128
4.3.5. Operatori compatti	135
4.3.6. L'estensione di operatori simmetrici	136
4.3.7. Teorema di Stone	138
4.3.8. L'integrale di Fourier	139
Appendice III	141
4.4. La formula di Lie-Trotter e la Meccanica Quantistica	141
4.5. Metodi computazionali	142
4.5.1. Implementazione numerica della formula di Lie-Trotter	142
4.5.2. Matrici sparse e problemi spettrali	144
4.5.3. L'algoritmo di Lanczos	145

Problemi	149
Soluzioni	157
Galleria di ritratti	167
IV - Applicazioni alla Teoria delle Probabilità	179
Capitolo 5. Caso e probabilità in Fisica	181
<i>Introduzione</i>	181
<i>Bibliografia ragionata</i>	182
5.1. La definizione matematica di probabilità	182
5.1.1. La formula di inclusione/esclusione	185
5.1.2. Probabilità condizionata	186
5.1.3. La formula di Bayes	187
5.1.4. Indipendenza statistica	187
5.2. Variabili aleatorie	187
5.2.1. Funzione di distribuzione	188
5.2.2. Funzione generatrice	188
5.2.3. Variabili indipendenti	189
5.2.4. Somma di variabili aleatorie	189
5.2.5. La legge dei grandi numeri	190
5.3. Processi aleatori	191
5.3.1. Catene di Markov discrete	191
5.3.2. Catene di Markov continue	195
5.3.3. Proprietà di convergenza	195
5.3.4. Il metodo MonteCarlo	196
5.3.5. Come realizzare successioni casuali	196
5.3.6. MonteCarlo dinamico	198
5.3.7. Equazione di Langevin	200
5.3.8. Equazione di Langevin e Meccanica Quantistica	201
Appendice IV	203
5.4. Distribuzioni di probabilità elementari	203
5.5. I teoremi di Gershgorin e Frobenius	204
5.6. Distribuzione di Polya	205
5.7. Soluzione esatta del modello di Ehrenfest	206
5.8. Calcolo del tempo medio di ritorno	207
5.9. Dimostrazione della Eq. (5.16)	208
5.10. Processi diffusivi	209
Problemi	213
Soluzioni	217
Bibliografia	225
Indice analitico	229
Elenco delle figure	235

Prefazione

La prima edizione di queste lezioni [Ono84] risale all'epoca "pre-computer", sia nel senso che allora non era ancora diffuso l'uso del calcolatore personale, e quindi l'approccio numerico a problemi complessi riguardava per lo più applicazioni avanzate, ma non faceva parte della formazione di base del Fisico, sia anche per il fatto che la preparazione del testo era ancora basata su manoscritto-dattilografia-stampa. Ciò ha comportato che, una volta esaurite le copie a suo tempo prodotte dall'Editore Zara, le Lezioni siano sopravvissute solo grazie alla disponibilità di qualche copia salvata dalla estinzione. Mi sono allora deciso, d'accordo con il titolare della Editrice Zara, cui vanno i miei ringraziamenti, a riportare le lezioni su supporto elettronico. Nell'occasione ho pensato che fosse tempo di rinnovare il testo aggiungendo quelle informazioni essenziali che riguardano l'applicazione di moderne tecniche numeriche alla soluzioni di problemi lineari, ormai entrate nell'uso molto diffuso tra Fisici a tutti i livelli. Il testo originale è quindi integrato con esercizi che richiedono l'uso del calcolo numerico; le soluzioni sono date utilizzando il linguaggio `matlab` in quanto si presta a scrivere un codice in modo molto leggibile, oppure nei casi dove è conveniente, il linguaggio simbolico `Mathematica`. Nel testo solo l'ossatura dei programmi sarà riportata, in modo da convogliare il nocciolo dell'idea, mentre i codici completi funzionanti sono riportati a parte in forma di archivio a sè stante, disponibile su rete. La copertina originale (una matrice iperbolica di dimensione 23) era stata a suo tempo prodotta "a mano"; ora è sufficiente utilizzare il codice `hpm.m` inserendo la prima riga della matrice. Le matrici iperboliche possono presentarsi in modo del tutto inaspettato, (si veda l'indice analitico) ponendo problemi interessanti. Per questo ho mantenuto la copertina originale, cui sono affezionato*.


Rispetto alla prima edizione, oltre all'aggiunta di esercizi e di esempi di applicazione di calcolo numerico/simbolico, ho inserito un capitolo sulle applicazioni alla teoria delle probabilità. Si tratta di un ambito in cui l'analisi lineare gioca un ruolo importante e l'argomento dovrebbe entrare, a mio giudizio, nel curriculum standard degli studenti di Fisica al I livello. Qui l'accento è posto più sulle tecniche di simulazione piuttosto che sull'analisi matematica. Il problema è costituito dal fatto che una presentazione matematicamente consistente prenderebbe facilmente troppo spazio rispetto a quanto previsto dalle ristrettezze di un corso semestrale e non si arriverebbe ad esplorare problemi interessanti. La possibilità di simulare numericamente situazioni anche complesse con il minimo sforzo permette invece di addentrarsi rapidamente nelle problematiche delle variabili aleatorie (*random variables*), delle catene di Markov, delle camminate aleatorie (*random walks*), dei processi diffusivi retti da equazioni differenziali stocastiche. Naturalmente le simulazioni non permettono mai di raggiungere risultati sicuri al 100%, ma forniscono in ogni caso una buona guida per congetturare il risultato cercato, e nella stragrande maggioranza dei

*Qualcuno appassionato di misteri alla Dan Brown (non è il caso dell'autore) potrebbe cogliere un messaggio premonitore nel fatto che la copertina, se vista da lontano, svela una **X** evidentemente legata, con ben 25 anni di anticipo, al "system X", attuale sistema operativo del Mac.

casi realistici è quanto abbiamo a disposizione (oltre all'esperimento reale) in mancanza di un risultato basato sull'analisi matematica.

Desidero ringraziare quelli tra i miei colleghi (in particolare *Raffaella Burioni* e *Gian Carlo Rossi*) che mi hanno onorato utilizzando la prima edizione del libro e mi hanno così stimolato a metterlo nella forma attuale in formato elettronico. Inoltre sono molto riconoscente a tutti i colleghi con cui ho goduto di una felice collaborazione in questi anni, *George E. Andrews, Marisa Bonini, Mario Casartelli, Marcello Ciafaloni, Giovanni Cicuta, Claudio Destri, Francesco Di Renzo, Roberto DePietri, Vladimir A. Fateev, Giorgio Giavarini, John R. Klauder, Giuseppe Marchesini, Pietro Menotti, Massimo Pauri, Giampiero Tecchiolli, Gabriele Veneziano*. Da ognuno di loro ho imparato qualcosa per me importante. Desidero inoltre ringraziare gli studenti che hanno frequentato il corso di *Metodi Probabilistici della Fisica* nel 2004–2005, per gli stimoli che mi hanno fornito rendendo il compito di docente piuttosto gradevole. Inoltre un grazie particolare all'amico *Augusto Schianchi*, collega della Facoltà di Economia, per la curiosità sempre vivace che dimostra per quei problemi di matematica che si collocano alla frontiera tra Economia e Fisica.

Notazioni

Il testo è articolato in capitoli, sezioni, problemi, soluzioni, osservazioni: queste ultime sono indicate con  (per invitare il lettore a rallentare un attimo e riflettere; il simbolo è scelto in omaggio a Donald E. Knuth e al suo manuale di T_EX). Le soluzioni dei problemi sono raggruppate alla fine di ciascun capitolo e identificate con $\#nn$. Un frammento di codice in un linguaggio quale `matlab` o `Mathematica` è segnalato da un formato particolare del tipo

```

      <nome del programma>
1      function Y=hpm(row)
2      .... % dettagli omessi
3      N=length(row);
4      U=eye(N,N);
5      X=diag(ones(N-1,1),-1);
6      X=X+X';
7      ... etc ...

```

mentre il codice completo si può recuperare in rete.

I teoremi sono di primaria importanza in Matematica, almeno da Euclide in poi. Per evidenziarli e facilitarne così la ricerca si è introdotta, in margine al testo, una etichetta che ne riporta il numero progressivo. Lo stesso si è fatto per le definizioni. I vettori sono normalmente indicati in grassetto, come ξ, η , etc.; ma preferibilmente utilizzeremo la notazione di Dirac $|\text{ket}\rangle$ e $\langle \text{bra}|$. Gli operatori lineari sono distinti dai caratteri $\mathcal{A}, \mathcal{B}, \dots, \mathcal{X}$; per i proiettori si riserva la notazione \mathbb{P}, \mathbb{E} , etc. La norma di vettori e operatori è indicata con $\|\xi\|$, $\|\mathcal{A}\|$, etc. Nella Parte IV si indica con $\mathbb{M}[\mathbf{x}]$ e $\mathbb{D}[\mathbf{x}]$ la media e la varianza della variabile aleatoria \mathbf{x} . I campi numerici sono contraddistinti come usuale da \mathbb{N} (numeri naturali), \mathbb{Z} (interi), \mathbb{R} (reali), \mathbb{C} (complessi), \mathbb{H} (quaternioni). Parte reale e immaginaria di un numero complesso z sono indicati con $\Re\{z\}$ e $\Im\{z\}$. Una relazione come $\mathcal{A} \leftarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ significa che l'operatore \mathcal{A} è *rappresentato* dalla matrice a destra; analogamente

Teor. x.y.z

Def. x.y.z

$|\alpha\rangle \leftrightarrow (\alpha_1 \ \alpha_2 \ \alpha_3 \ \dots)$ assegna al vettore astratto $|\alpha\rangle$ la sua rappresentazione in una specifica base. La trasposta di una matrice A è indicata (come in `matlab`) con A' . Il cartello $n \neq \infty$ sarà utilizzato per segnalare argomenti dove la dimensione finita dello spazio lineare è cruciale e non è possibile estrapolare in modo ingenuo a spazi ∞ -dimensionali oppure che si tratta di proprietà specifiche dello spazio di Hilbert che non hanno analogo in un numero finito di dimensioni. I teoremi sono numerati in base al capitolo. Il termine di una dimostrazione è segnalato da ■, quello di un esempio o di un commento che interrompono il testo da ◀. Gran parte dei problemi sono raggruppati alla fine di ciascuna parte, alcuni sono invece distribuiti nel testo (voce “problemi” nell’indice analitico) allo scopo di stimolare il lettore a verificare il suo grado di comprensione.

I riferimenti bibliografici sono etichettati con la convenzione `amsalpha`, es. [Dir29] abbreviazione di *Dirac 1929*. Il testo è composto nel font *AMS_Euler* disegnato espressamente da Donald Knuth per avvicinarsi agli standard dei Matematici. Un fondamentale aiuto mi è arrivato dall’ottima introduzione a L^AT_EX in [GMS94], ora in edizione rinnovata (2004).

Prefazione alla prima edizione

Queste lezioni di Analisi lineare rappresentano una parte del programma del corso di Metodi Matematici della Fisica che ho svolto per qualche anno presso l’Università degli Studi di Parma.

.....

Il carattere del corso intende realizzare un compromesso tra la mole di argomenti trattati e il livello di rigore e di completezza. Mentre la parte relativa agli spazi ad un numero finito di dimensioni è sviluppata in un modo che ritengo coerente e sufficientemente generale, quella relativa agli spazi di Hilbert soffre, di necessità, di ampie lacune ed è inoltre sviluppata ad un livello per lo più euristico. Secondo le mie intenzioni, lo studente attento deve sentirsi insoddisfatto di questa parte del programma ed essere quindi spronato ad approfondirla sui vari trattati che vengono consigliati nella bibliografia. Ho preferito impostare il corso su un atteggiamento pragmatico (risolvere equazioni integrali senza sapere necessariamente che si chiamano di Fredholm) ed ho quindi privilegiato la parte di esercizi rispetto ad un maggiore sviluppo della teoria formale. L’obiettivo ideale è quello di convincere lo studente che è importante avere a disposizione una vasta varietà di strumenti matematici in modo da potere scegliere quello più adatto per una particolare applicazione. Un limite del corso, a questo riguardo, è semmai il suo carattere astratto - problemi di calcolo effettivo, basati sulla analisi numerica, richiederebbero un corso a parte.[†]

Due parole sulla impostazione di queste lezioni:

- *Parte I - È una introduzione generale che vuole giustificare lo studio dell’analisi lineare con esempi tratti dalla Fisica classica. Sarebbe opportuno, alla fine del corso, ritornare a questa prima parte e riesaminare quei punti che sono difficilmente comprensibili in una prima lettura.*
- *Parte II - Spazi finito-dimensionali. Il problema centrale è quello della decomposizione spettrale di una qualunque operatore lineare. Ho scelto di fondare la teoria*

[†]In questa seconda edizione si è infatti cercato di completare il testo in questa direzione.

sulla formula di Riesz-Dunford, che richiede alcune nozioni di funzioni analitiche ed obbliga perciò a un certo lavoro di sintesi con l'altra parte del corso di M.M.

- *Parte III - Spazio di Hilbert.* Ho preferito sorvolare sui problemi della teoria della misura (Lebesgue) in favore di un maggiore spazio dedicato agli “spazi funzionali” (spazi di funzioni analitiche) che permettono di sviluppare numerosi esempi elementari e sono anche importanti per le applicazioni (stati coerenti in meccanica quantistica). La teoria spettrale, basata sull'integrale di Cauchy-Dunford, si riallaccia in modo euristico, ma efficace, al caso finito-dimensionale. [‡].

Sono felice di ringraziare in questa sede tutti i miei colleghi che mi hanno in qualche modo aiutato a preparare queste lezioni,

Mario Casartelli, Giuseppe Marchesini e Massimo Pauri

in particolare per tutto quello che mi hanno insegnato in tanti anni di lavoro in comune. Ed infine un grazie di cuore a mia moglie Liliana per l'accurato lavoro di dattilografia (gli errori di battitura sono opera mia).

Parma, giugno 1984.

[‡]...e ora anche una Parte IV

I - Introduzione

CAPITOLO 1

Spazi lineari in Meccanica Classica

1.1. Il tensore d'inerzia

Consideriamo un corpo rigido \mathfrak{R} con un punto fisso \mathcal{O} . Il momento angolare del corpo rispetto al punto \mathcal{O} è dato da

$$\mathbf{M} = \int_{\mathfrak{R}} \rho(\mathbf{r}) \mathbf{r} \wedge \frac{d\mathbf{r}}{dt} dV$$

avendo fissato la seguente notazione: ρ è la densità di massa al punto \mathbf{r} ; dV è l'elemento di volume. Essendo per un corpo rigido $d\mathbf{r}/dt = \boldsymbol{\omega} \wedge \mathbf{r}$, si ottiene *

$$\begin{aligned} \mathbf{M} &= \int_{\mathfrak{R}} \rho(\mathbf{r}) \mathbf{r} \wedge (\boldsymbol{\omega} \wedge \mathbf{r}) dV \\ &= \int_{\mathfrak{R}} \rho(\mathbf{r}) ((\mathbf{r} \cdot \mathbf{r}) \boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \mathbf{r}) \mathbf{r}) dV \end{aligned}$$

In componenti

$$\begin{aligned} M_i &= \int_{\mathfrak{R}} \rho(\mathbf{r}) ((\mathbf{r} \cdot \mathbf{r}) \omega_i - (\boldsymbol{\omega} \cdot \mathbf{r}) x_i) dV \\ &= \sum_{j=1}^3 \int_{\mathfrak{R}} \rho(\mathbf{r}) (r^2 \delta_{ij} - x_i x_j) \omega_j dV \equiv \sum_j \mathcal{I}_{ij} \omega_j. \end{aligned}$$

Le nove quantità $\mathcal{I}_{ij} = \int_{\mathfrak{R}} \rho(\mathbf{r}) (r^2 \delta_{ij} - x_i x_j) dV$ formano una matrice quadrata (tensore d'inerzia). La relazione tra \mathbf{M} e $\boldsymbol{\omega}$ non è dunque di proporzionalità, in generale. Tuttavia \mathbf{M} e $\boldsymbol{\omega}$ sono legati da una *relazione lineare*, il che significa che indicando con $\mathbf{M}(\boldsymbol{\omega})$ la funzione $\boldsymbol{\omega} \rightarrow \mathbf{M}$ si ha:

$$\begin{aligned} \mathbf{M}(\lambda \boldsymbol{\omega}) &= \lambda \mathbf{M}(\boldsymbol{\omega}) \\ \mathbf{M}(\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) &= \mathbf{M}(\boldsymbol{\omega}_1) + \mathbf{M}(\boldsymbol{\omega}_2). \end{aligned}$$

Si indica perciò $\mathbf{M} = \mathcal{I} \boldsymbol{\omega}$ e si dice che \mathcal{I} è un **operatore lineare**, di cui $\{\mathcal{I}_{ij}\}$ è la *rappresentazione matriciale* nella base delle coordinate cartesiane individuata dai versori $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$. Anche se in generale \mathbf{M} e $\boldsymbol{\omega}$ non sono paralleli, ci si può chiedere se tale parallelismo non possa verificarsi per qualche particolare direzione di $\boldsymbol{\omega}$. Ciò porta ad impostare l'equazione

$$\mathcal{I} \boldsymbol{\omega} = \lambda \boldsymbol{\omega}.$$

*Ricordiamo che per tre vettori qualunque $\mathbf{a}, \mathbf{b}, \mathbf{c}$, vale la formula $\mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

In componenti, si ottiene

$$\sum_{j=1}^3 \mathcal{I}_{ij} \omega_j = \lambda \omega_i$$

$$\sum_{j=1}^3 (\mathcal{I}_{ij} - \lambda \delta_{ij}) \omega_j = 0.$$

Si tratta dunque di risolvere un sistema lineare omogeneo di tre equazioni in tre incognite. Tale sistema ammette l'unica soluzione $\boldsymbol{\omega} = 0$ (che non ci interessa) a meno che non sia

$$(1.1) \quad \det |\mathcal{I}_{ij} - \lambda \delta_{ij}| = 0.$$

Prima di considerare esplicitamente quest'ultima equazione, è possibile anticipare alcune proprietà delle soluzioni. È importante riconoscere che la matrice \mathcal{I} è *simmetrica*, cioè $\mathcal{I}_{ij} = \mathcal{I}_{ji}$. Questo fatto ha notevoli conseguenze:

i) il numero λ deve essere reale, infatti

$$\lambda = \frac{\sum_{ij} \mathcal{I}_{ij} \omega_i \omega_j}{\sum_i |\omega_i|^2} = \bar{\lambda}$$

ii) Se λ_1 e λ_2 sono due valori distinti per cui il sistema lineare ammette soluzioni $\boldsymbol{\omega}^{(1)}$ e $\boldsymbol{\omega}^{(2)}$ entrambe non nulle, allora $\boldsymbol{\omega}_1 \cdot \boldsymbol{\omega}_2 = 0$.

Infatti sia

$$\sum_j \mathcal{I}_{ij} \omega_j^{(1)} = \lambda_1 \omega_i^{(1)}$$

$$\sum_j \mathcal{I}_{ij} \omega_j^{(2)} = \lambda_2 \omega_i^{(2)}$$

Moltiplichiamo la prima equazione per $\omega_i^{(2)}$ e sommiamo sull'indice i ; moltiplichiamo la seconda equazione per $\omega_i^{(1)}$ e sommiamo su i . Otteniamo allora

$$\sum_{ij} \mathcal{I}_{ij} \omega_i^{(2)} \omega_j^{(1)} = \lambda_1 \boldsymbol{\omega}^{(2)} \cdot \boldsymbol{\omega}^{(1)}$$

$$\sum_{ij} \mathcal{I}_{ij} \omega_i^{(1)} \omega_j^{(2)} = \lambda_2 \boldsymbol{\omega}^{(1)} \cdot \boldsymbol{\omega}^{(2)}$$

I due membri a sinistra sono uguali per via della simmetria di \mathcal{I} . Dunque

$$(\lambda_1 - \lambda_2) \boldsymbol{\omega}^{(1)} \cdot \boldsymbol{\omega}^{(2)} = 0$$

ed essendo $\lambda_1 \neq \lambda_2$, si ottiene che $\boldsymbol{\omega}^{(1)}$ e $\boldsymbol{\omega}^{(2)}$ sono ortogonali. Consideriamo dunque l'equazione (1.1), e cioè

$$\begin{vmatrix} \mathcal{I}_{11} - \lambda & \mathcal{I}_{12} & \mathcal{I}_{13} \\ \mathcal{I}_{21} & \mathcal{I}_{22} - \lambda & \mathcal{I}_{23} \\ \mathcal{I}_{31} & \mathcal{I}_{32} & \mathcal{I}_{33} - \lambda \end{vmatrix} = 0$$

Sviluppando il determinante e raccogliendo i termini omogenei in λ :

$$\lambda^3 - p_1 \lambda^2 + p_2 \lambda - p_3 = 0$$

dove

$$\begin{aligned} p_1 &= \sum_i \mathcal{I}_{ii} \\ p_2 &= \sum_{i < j} \mathcal{I} \begin{pmatrix} i & j \\ i & j \end{pmatrix} \equiv \begin{vmatrix} \mathcal{I}_{11} & \mathcal{I}_{12} \\ \mathcal{I}_{21} & \mathcal{I}_{22} \end{vmatrix} + \begin{vmatrix} \mathcal{I}_{11} & \mathcal{I}_{13} \\ \mathcal{I}_{31} & \mathcal{I}_{33} \end{vmatrix} + \begin{vmatrix} \mathcal{I}_{22} & \mathcal{I}_{23} \\ \mathcal{I}_{32} & \mathcal{I}_{33} \end{vmatrix} \\ p_3 &= \sum_{i < j < k} \mathcal{I} \begin{pmatrix} i & j & k \\ i & j & k \end{pmatrix} \equiv \det |\mathcal{I}| \end{aligned}$$

(si veda in generale l'Eserc. 30, Parte II, formule di Newton).

Abbiamo visto che l'equazione (1.1) ammette solo radici reali; di più tali radici sono positive dato che $\boldsymbol{\omega} \cdot \mathcal{I} \boldsymbol{\omega} = \lambda(\boldsymbol{\omega} \cdot \boldsymbol{\omega})$ e il membro di sinistra è positivo, dato che vale $2T$ (due volte l'energia cinetica). Sia $\boldsymbol{\omega}^{(1)}$ una soluzione del sistema lineare $\mathcal{I} \boldsymbol{\omega} = \lambda_1 \boldsymbol{\omega}$. Supponiamo di ruotare gli assi delle coordinate in modo che uno degli assi sia nella direzione di $\boldsymbol{\omega}^{(1)}$. In questo nuovo sistema di coordinate la matrice che rappresenta \mathcal{I} avrà la seguente espressione

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \mathcal{I}'_{22} & \mathcal{I}'_{23} \\ 0 & \mathcal{I}'_{32} & \mathcal{I}'_{33} \end{pmatrix}$$

La ricerca delle altre soluzioni si può quindi ridurre al piano ortogonale a $\boldsymbol{\omega}^{(1)}$. L'Eq. (1.1) in questa nuova base è perciò

$$(\lambda - \lambda_1) (\lambda^2 - (\mathcal{I}'_{22} + \mathcal{I}'_{33}) \lambda + \mathcal{I}'_{22} \mathcal{I}'_{33} - \mathcal{I}'_{23} \mathcal{I}'_{32}) = 0$$

che, oltre a λ_1 ammette le soluzioni

$$\lambda_{2,3} = \frac{1}{2} \left(\mathcal{I}'_{22} + \mathcal{I}'_{33} \pm \sqrt{(\mathcal{I}'_{22} - \mathcal{I}'_{33})^2 + 4\mathcal{I}'_{23}^2} \right)$$

Come si vede le soluzioni sono sempre reali e positive (in particolare si può avere $\lambda_2 = \lambda_3$ solo se $\mathcal{I}'_{22} = \mathcal{I}'_{33}$ e $\mathcal{I}'_{23} = 0$). Ricordiamo che le direzioni individuate dai vettori $\boldsymbol{\omega}^{(1)}$, $\boldsymbol{\omega}^{(2)}$ e $\boldsymbol{\omega}^{(3)}$ ortogonali si dicono *assi principali* del corpo rigido. Tali assi sono individuati univocamente solo se $\lambda_1 \neq \lambda_2 \neq \lambda_3$. Infatti se capita che ad es. $\lambda_2 = \lambda_3$ allora ogni combinazione lineare

$$\boldsymbol{\omega} = \mathbf{a} \boldsymbol{\omega}^{(2)} + \mathbf{b} \boldsymbol{\omega}^{(3)}$$

soddisfa l'equazione $\mathcal{I} \boldsymbol{\omega} = \lambda_2 \boldsymbol{\omega}$. Notiamo che questo fatto si verifica se il corpo ha simmetria cilindrica, ma non vale il viceversa. Ad es. il tensore d'inerzia di un cubo omogeneo ha tutti i momenti d'inerzia coincidenti, pur non avendo il cubo simmetria sferica. Agli effetti inerziali tuttavia il cubo si comporta come una sfera.

1.2. Modi normali di due oscillatori accoppiati

Quale secondo esempio, consideriamo un semplice sistema meccanico costituito da due masse uguali vincolate a scorrere senza attrito lungo una retta; su di esse agiscono forze elastiche come indicato nella Fig. 1-1

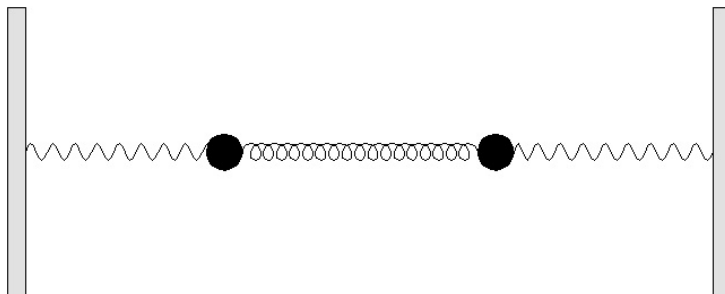


FIGURA 1-1. Oscillatori accoppiati

Le due molle fissate alle pareti abbiano modulo di elasticità k e la molla che accoppia le masse abbia modulo λ . Indicando con x_1 e x_2 gli scostamenti dalle posizioni di equilibrio delle due masse, l'energia potenziale del sistema è data da

$$\begin{aligned} V(x_1, x_2) &= \frac{1}{2} k x_1^2 + \frac{1}{2} k x_2^2 + \frac{1}{2} \lambda (x_1 - x_2)^2 \\ &= \frac{1}{2} (k + \lambda) (x_1^2 + x_2^2) - \lambda x_1 x_2 \end{aligned}$$

Adottando unità di misura in cui la massa sia uno, le equazioni del moto sono

$$(1.2) \quad \begin{aligned} \ddot{x}_1 &= -\frac{\partial V}{\partial x_1} = -(k + \lambda) x_1 + \lambda x_2 \\ \ddot{x}_2 &= -\frac{\partial V}{\partial x_2} = \lambda x_1 - (k + \lambda) x_2 \end{aligned}$$

Un modo elementare per risolvere questo sistema di equazioni lineari è dato dalla sostituzione

$$(1.3) \quad \xi = (x_1 + x_2)/\sqrt{2}; \quad \eta = (x_2 - x_1)/\sqrt{2}$$

Si trova infatti

$$\begin{aligned} \ddot{\xi} &= -k \xi \\ \ddot{\eta} &= -(k + 2\lambda) \eta \end{aligned}$$

Le equazioni sono ora separate e la soluzione è data da

$$\xi(t) = A \cos(\omega t + \delta); \quad \eta(t) = B \cos(\omega' t + \epsilon);$$

con $\omega = \sqrt{k}$ e $\omega' = \sqrt{k + 2\lambda}$; A, B, δ, ϵ sono costanti arbitrarie. Il sistema di equazioni può essere scritto in forma più compatta introducendo un vettore a due componenti $\mathbf{x} = (x_1, x_2)$ e l'operatore lineare Ω^2 definito da

$$\Omega^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} k + \lambda & -\lambda \\ -\lambda & k + \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

per cui

$$\frac{d^2 \mathbf{x}}{dt^2} = -\Omega^2 \mathbf{x}.$$

È immediato riconoscere che la trasformazione (1.3) altro non è che una rotazione del piano. Si può dunque affermare che il sistema è stato risolto introducendo quel sistema di riferimento in cui la matrice Ω^2 assume la forma più semplice, cioè quella diagonale. Vedremo d'altronde che per ogni matrice simmetrica (qual'è Ω^2) *esiste una rotazione che la trasforma nella forma diagonale*. La trasformazione Eq. (1.3) può essere trovata cercando le direzioni per cui vale l'equazione

$$\Omega^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \omega^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

in modo analogo a quanto si è fatto per individuare gli assi principali del corpo rigido. In generale un sistema lineare del tipo

$$\frac{d^2}{dt^2} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = -\Omega^2 \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

può essere separato in n equazioni disaccoppiate mediante una rotazione nello spazio (x_1, x_2, \dots, x_n) , purché la matrice Ω^2 sia simmetrica. Un esempio è offerto nella sezione seguente.

🔗 *La soluzione del problema precedente è elementare e certo non richiede l'ausilio del calcolatore. Tuttavia si presta a evidenziare alcune tecniche di programmazione **matlab** che sono utili in casi più complessi. Dovendo risolvere il sistema (1.2) si può ricorrere a una delle routines di soluzione di equazioni differenziali che le librerie di **matlab** mettono a disposizione. Le equazioni risolubili con **ode45** (o con le altre routines simili) sono del tipo*

$$\frac{dy(t)}{dt} = f(t, y(t))$$

dove $y(t)$ può essere un vettore di qualunque dimensione. Se scegliamo

$$y(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix}$$

la soluzione si può impostare così (ponendo $m = 1$)

```

                                dueosc.m
function ydot = dueosc(t,x,k,lam)
x1=x(1); x2=x(2); v1=x(3); v2=x(4);
ydot = [          v1;          v2;
        -(k+lam)*x1+lam*x2; lam*x1 - (k+lam)*x2];

```

e una semplice chiamata del tipo

```

[T,X] = ode45(@dueosc, [0,100], [1,0,0,1], [], 1, .1);

```

ritorna un vettore di tempi $T = \{t_1 = 0, t_2, \dots, t_N = 100\}$ e una matrice

$$X = \begin{bmatrix} x_1(t_1) & x_2(t_1) & v_1(t_1) & v_2(t_1) \\ x_1(t_2) & x_2(t_2) & v_1(t_2) & v_2(t_2) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(t_N) & x_2(t_N) & v_1(t_N) & v_2(t_N) \end{bmatrix}$$

che contiene la soluzione “campionata” negli istanti $\{t_j\}$. Ovviamente questo modo di risoluzione si applica anche nel caso di equazioni non lineari. ◀

1.3. Modi normali di n oscillatori accoppiati

La trattazione del sistema descritto nella sezione precedente può essere generalizzata al caso di n masse interagenti tramite forze elastiche. Esamineremo ora questo problema in quanto mostra chiaramente l'utilità del calcolo matriciale. Il concetto di “modo normale” è pure molto importante per le applicazioni. Consideriamo dunque una molla di cui trascuriamo la massa; ad essa sono attaccate a uguale distanza n masse identiche; gli estremi della molla sono fissi e si trascurano tutti gli attriti (vedi Fig. 1-2).

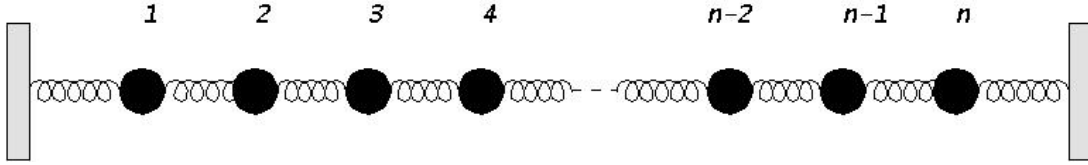


FIGURA 1-2. n oscillatori accoppiati.

Supponiamo che le masse possano subire solo spostamenti longitudinali. Il moto trasversale si può studiare in modo analogo. Siano x_1, x_2, \dots, x_n gli spostamenti delle masse dai punti di equilibrio. L'energia potenziale del sistema sarà allora data da

$$V = \frac{1}{2} \kappa \sum_{j=1}^{n+1} (x_j - x_{j-1})^2$$

dove $\kappa/(n+1)$ è la costante di elasticità della molla e si conviene che $x_0 = x_{n+1} = 0$. Tenendo conto che l'energia cinetica è la somma $\frac{1}{2} m \sum_j \dot{x}_i^2$, le equazioni del moto sono date da

$$m \ddot{x}_j = \kappa (x_{j-1} - 2x_j + x_{j+1})$$

e possono essere scritte in forma matriciale introducendo la matrice colonna $\|x_j\|$:

$$(1.4) \quad m \frac{d^2}{dt^2} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_\ell \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = -\kappa \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \dots & -1 & 2 & -1 & \dots & 0 \\ \vdots & \vdots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & -1 & 2 & -1 \\ 0 & 0 & \dots & 0 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_\ell \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}$$

ovvero

$$\ddot{\mathbf{x}} = -\Omega^2 \mathbf{x}$$

La matrice indicata con Ω^2 è simmetrica: come vedremo in seguito (Cap.2) esiste allora una rotazione di assi \mathcal{R} nello spazio x_1, x_2, \dots, x_n tale che Ω^2 diventa una matrice diagonale; le variabili dinamiche $\xi_i = \sum \mathcal{R}_{ij} x_j$ si dicono “modi normali” del sistema ed hanno una evoluzione temporale periodica

$$\xi_i(t) = \xi_i(0) \cos \omega_i t + \frac{\dot{\xi}_i(0)}{\omega_i} \sin \omega_i t.$$

Vediamo come si trova esplicitamente la rotazione \mathcal{R} . Si tratta in pratica di cercare le soluzioni del sistema lineare

$$\ddot{x}_i = -\omega^2 x_i \Rightarrow \sum_j (\omega^2 \delta_{ij} - \Omega_{ij}^2) x_j = 0.$$

Il sistema ammette soluzioni se il determinante

$$(1.5) \quad \begin{vmatrix} 2-\lambda & -1 & 0 & 0 & \dots & 0 \\ -1 & 2-\lambda & -1 & 0 & \dots & 0 \\ 0 & -1 & 2-\lambda & -1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -1 & 2-\lambda & -1 \\ 0 & 0 & \dots & 0 & -1 & 2-\lambda \end{vmatrix}$$

si annulla. Otteniamo così un’equazione algebrica di grado n nella variabile $\lambda = \omega^2$ e questa ammette sempre n radici complesse; ma in questo caso le radici sono tutte reali e positive (di ciò daremo una dimostrazione in generale nel Cap.2). Una giustificazione intuitiva è fornita dal fatto che radici negative o complesse corrisponderebbero a un moto smorzato, mentre abbiamo escluso la presenza di forze dissipative, oppure con ampiezza crescente che contravverrebbero alla conservazione dell’energia totale.

Il metodo più generale per calcolare il determinante Eq. (1.5) è il seguente: sviluppiamo il determinante secondo la prima riga e otteniamo

$$\det(\Omega^2 - \lambda \mathbb{I}) \equiv \Delta_n = (2 - \lambda) \Delta_{n-1} - \Delta_{n-2}$$

avendo indicato con $\Delta_{n-1}, \Delta_{n-2}, \dots$ determinanti simili a quello che ci interessa ma con dimensione via via decrescente. L’equazione

$$\Delta_n + (\lambda - 2) \Delta_{n-1} + \Delta_{n-2} = 0$$

è una *equazione alle differenze finite*, lineare e omogenea a coefficienti costanti (rispetto a n). La soluzione generale è data dalla combinazione lineare di due soluzioni indipendenti, le quali si trovano (in analogia con il caso delle equazioni differenziali) ponendo $\Delta_n = e^{n\alpha}$ che ci dà

$$e^{n\alpha} + (\lambda - 2) e^{(n-1)\alpha} + e^{(n-2)\alpha} = 0$$

ossia

$$e^{2\alpha} + (\lambda - 2) e^{\alpha} + 1 = 0$$

da cui

$$z = e^{\alpha} = \frac{1}{2}(2 - \lambda) \pm \frac{1}{2}\sqrt{(\lambda - 2)^2 - 4}$$

Posto $\lambda = 2(1 - \cos \vartheta)$ si trova $z = e^{\pm i\vartheta}$ e dunque la soluzione generale si trova nella forma

$$\Delta_n = A e^{in\vartheta} + B e^{-in\vartheta}$$

Imponendo le condizioni “al contorno” (c.c.) ($n = 1, 2$) si determinano A e B :

$$n = 1 \implies A e^{i\vartheta} + B e^{-i\vartheta} = 2 - \lambda = 2 \cos \vartheta$$

$$n = 2 \implies A e^{2i\vartheta} + B e^{-2i\vartheta} = (2 - \lambda)^2 - 1 = 4 \cos^2 \vartheta - 1$$

Con un po' di algebra si trova $A = \bar{B} = -\frac{1}{2} i e^{i\vartheta} / \sin \vartheta$ e infine

$$\Delta_n = \frac{\sin((n+1)\vartheta)}{\sin \vartheta}$$

Gli zeri del determinante, ossia le frequenze proprie di vibrazione del sistema elastico, sono date perciò da

$$\omega_j = \sqrt{\lambda_j} = 2\sqrt{k/m} \sin\left(\frac{j\pi/2}{n+1}\right), j = 1, 2, \dots, n.$$

Indichiamo con $\xi^{(i)}$ le soluzioni dell'equazione

$$\Omega^2 \xi^{(i)} = \omega_j^2 \xi^{(i)}$$

cioè i modi normali. È possibile fissare un fattore moltiplicativo arbitrario in modo tale che

$$\sum_{k=1}^n |\xi_k^{(j)}|^2 = 1.$$

La matrice $\mathcal{R}_{jk} = \xi_k^{(j)}$ rappresenta la rotazione che porta Ω alla forma diagonale. Esplicitamente si trova

$$\xi_k^{(j)} = \sqrt{\frac{2}{n+1}} \sin \frac{kj\pi}{n+1}.$$

Lo spettro di frequenze, nel caso di valori molto grandi di n , diventa essenzialmente quello di una corda di violino - cioè frequenze multiple di una frequenza fondamentale $\omega_1 \approx 2\sqrt{k/m} \frac{\pi}{n+1}$.

🔍 Si noterà che i modi normali $\xi_j^{(n)}$ si annullano per $j = 0$ e $j = n+1$, il che corrisponde alla convenzione $x_0 = x_{n+1} = 0$. In effetti un metodo più veloce per determinare le soluzioni consiste nel ricercare soluzioni del tipo $\xi_j = e^{i\alpha j}$ e nell'imporre la condizione al contorno alla soluzione generale. Tuttavia il metodo illustrato qui è più potente. Nel caso di masse diverse oppure molle di costante elastica differente, in ogni caso il calcolo del determinante si può impostare in modo ricorsivo, e risolvere in modo efficiente, almeno per via numerica. Ogni qual volta un problema complesso è ricondotto ad una relazione di ricorrenza che può essere risolta in un numero di passi proporzionale al numero di gradi di libertà il problema si può considerare risolto, da un punto di vista pratico. Osservare ad esempio il grande risparmio di operazioni aritmetiche nel calcolo attraverso la relazione di ricorrenza rispetto al calcolo di un determinante generico ($n!$ operazioni). ◀

🔍 Anche per il problema precedente sarebbe superfluo ricorrere al calcolatore. Tuttavia se le forze tra le masse fossero leggermente diverse, magari con correzioni al di là della approssimazione elastica, e conseguentemente le equazioni del moto fossero non-lineari, il ricorso al calcolatore risulta indispensabile. Questa è stata una delle prime applicazioni dei calcolatori elettronici, grazie a una felice idea di Enrico Fermi. Per saperne di più provare a chiedere a Google “Fermi Pasta Ulam”. Rimanendo nel contesto lineare, l'analisi numerica può fornire tecniche per il

calcolo dello spettro di matrici anche di grandi dimensioni. Nel caso degli n oscillatori accoppiati (Eq. (1.4)) lo spettro di frequenze si trova con

```

                                n o s c . m
F = diag(ones(n,1)) - ...      % diagonale principale
    diag(ones(n-1,1),-1);      % diagonale (i-1,i)
F = F + F';                    % simmetrizzare F
frequenze = sqrt(eig(F));      % eig ritorna lo spettro di F

```

avendo posto $k = m = 1$ senza perdere in generalità. F è una matrice tridiagonale, dunque una tipica matrice sparsa. In questi casi, se è necessario studiare matrici di grandi dimensioni, si possono utilizzare matrici specialmente studiate per questi casi. Ad es.

```

                                s n o s c . m - v e r s i o n e  ' ' s p a r s a ' '
F = spdiag(ones(n,1),0,n,n) ... % diagonale principale
    -spdiags(ones(n-1,1),-1,n,n); % diagonale (i-1,i)
F = F + F';                      % simmetrizzare F
freq = sqrt(eigs(F),K,'SA'); % eigs ritorna le prime "K" frequenze

```

La routine `eigs` permette di studiare lo spettro di matrici sparse di grandi dimensioni - e per grandi si intende in certi casi davvero *grandi*, $10^6 \times 10^6$ o più. ◀

1.4. Spazi a infinite dimensioni

1.4.1. Un sistema con n particelle e $n \rightarrow \infty$. Vorrei illustrare un problema fisico in cui entra la considerazione di uno spazio lineare infinito dimensionale. Consideriamo un sistema costituito da una corda elastica (lunga L , fissa agli estremi e senza massa) a cui sono attaccate n masse uguali a distanza $L/(n+1)$ libere di oscillare in senso trasversale. La Lagrangiana del sistema (analogamente al caso longitudinale) è data da

$$\mathcal{L} = \frac{1}{2}m \sum_{q=1}^n \dot{y}_q^2 - \frac{1}{2}\kappa \sum_{q=1}^{n+1} (y_q - y_{q-1})^2.$$

Indichiamo con $M = nm$ la massa totale, con $x = qa$ la posizione della q -esima particella, e $\kappa = KL/a$. Abbiamo visto che il moto è equivalente a quello di n oscillatori indipendenti (modi normali). Precisamente

$$\begin{aligned}
 y_q &= \sqrt{\frac{2}{n+1}} \sum_{p=1}^n \sin\left(\frac{pq\pi}{n+1}\right) \eta_p \\
 \eta_p &= \sqrt{\frac{2}{n+1}} \sum_{q=1}^n \sin\left(\frac{pq\pi}{n+1}\right) y_q \\
 \ddot{\eta}_p &= -\omega_p^2 \eta_p \\
 \omega_p &= 2\sqrt{\kappa/m} \sin\left(\frac{p\pi}{2(n+1)}\right).
 \end{aligned}$$

Supponiamo di considerare un numero n molto grande di particelle mantenendo però fissata la massa totale M . Al limite $n \rightarrow \infty$ il sistema può essere semplicemente descritto da una funzione $y(x) = y_{[x/a]}$ considerando x come variabile continua ($[x/a]$ indica la

parte intera di x/a). Si avrà

$$\begin{aligned}
 (1.6) \quad \mathcal{L} &= \frac{1}{2} \frac{M}{L} \sum_{q=1}^n \dot{y}(q a)^2 - \frac{1}{2} K L \sum_{q=1}^{n+1} a \left(\frac{y(q a) - y(q a - a)}{a} \right)^2 \\
 &\approx \frac{1}{2} \frac{M}{L} \int_0^L \left(\frac{\partial y}{\partial t} \right)^2 dx - \frac{1}{2} K L \int_0^L \left(\frac{\partial y}{\partial x} \right)^2 dx \\
 &= \int_0^L \left\{ \frac{1}{2} \rho \left(\frac{\partial y}{\partial t} \right)^2 - \frac{1}{2} \tau \left(\frac{\partial y}{\partial x} \right)^2 \right\} dx
 \end{aligned}$$

dove $\rho = M/L$ è la densità di massa lineare e $\tau = KL$ è la tensione della corda. L'equazione del moto diventa

$$\frac{\partial^2 y}{\partial t^2} = \frac{KL^2}{M} \frac{y(q a + a) - 2y(q a) + y(q a - a)}{a^2}$$

e al limite per $a \rightarrow 0$, considerando soluzioni "lisce" ossia continue e differenziabili,

$$\frac{\partial^2 y}{\partial t^2} = \frac{\tau}{\rho} \frac{\partial^2 y}{\partial x^2}$$

nota come l'equazione d'onda. I modi normali assumono una forma semplice al limite:

$$\begin{aligned}
 y(x) &= \sqrt{\frac{2}{(n+1)a}} \sum_{p=1}^n \sin\left(\frac{p\pi x}{L}\right) \sqrt{a} \eta_p \\
 &\downarrow (n \rightarrow \infty) \\
 y(x) &= \sqrt{\frac{2}{L}} \sum_{p=1}^{\infty} \sin\left(\frac{p\pi x}{L}\right) \phi_p.
 \end{aligned}$$

Con una serie di trasformazioni si può esprimere tutte le grandezze che entrano nella dinamica del sistema in termini continui:

$$\begin{aligned}
 \mathcal{L} &= \frac{1}{2} \frac{Ma}{L} \sum_{p=1}^n \dot{\eta}_p^2 - \frac{2KL}{a} \sum_{p=1}^n \sin^2\left(\frac{p\pi}{2(n+1)}\right) \eta_p^2 \\
 &\rightarrow \frac{1}{2} \rho \sum_{p=1}^{\infty} (\dot{\phi}_p^2 - \omega_p^2 \phi_p^2) \\
 (\omega_p &= \sqrt{\frac{K}{M}} p\pi, p = 1, 2, 3, \dots).
 \end{aligned}$$

Le variabili ϕ_p sono esprimibili in funzione di $y(x)$ secondo le relazioni

$$\begin{aligned}
 \phi_p &= \sqrt{a} \eta_p = \sqrt{\frac{2a}{n+1}} \sum_{q=1}^n \sin\left(\frac{p q \pi}{n+1}\right) y_q \\
 &= \sqrt{2/L} \sum_{q=1}^n \sin\left(\frac{p \pi x}{L}\right) y(x) a \\
 &\rightarrow \sqrt{2/L} \int_0^L \sin\left(\frac{p \pi x}{L}\right) y(x) dx
 \end{aligned}$$

Abbiamo così ottenuto un limite continuo per $n \rightarrow \infty$. La situazione di può riassumere in questi termini. La descrizione matematica di una fune elastica con massa uniformemente distribuita è fornita dalla funzione $y(x, t)$ che rappresenta a t fissato la forma della fune.

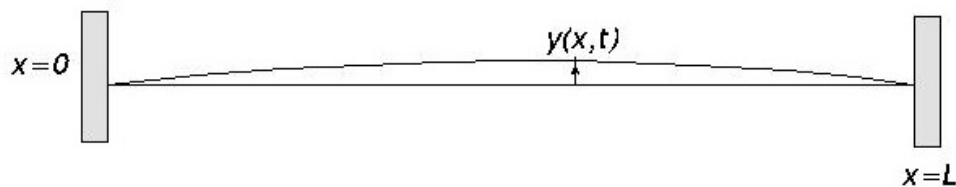


FIGURA 1-3. La corda vibrante.

La Lagrangiana del sistema è data dalla Eq.(1.6) da cui si ricava l'equazione del moto attraverso il principio variazionale di Euler-Lagrange (noi qui le abbiamo ottenute come limite del caso discreto). Il sistema presenta moti armonici collettivi analoghi ai modi normali di una catena discreta, in particolare si ha

$$y(x) = \sqrt{2/L} \sum_{p=1}^{\infty} \sin\left(\frac{p\pi x}{L}\right) \phi_p$$

$$\phi_p = \sqrt{2/L} \int_0^L \sin\left(\frac{p\pi x}{L}\right) y(x) dx$$

$$\mathcal{L} = \frac{1}{2} \rho \sum_{p=1}^{\infty} (\dot{\phi}_p^2 - \omega_p^2 \phi_p^2), \quad \omega_p = \sqrt{\frac{K}{M}} p \pi.$$

La trasformazione $y(x) \Leftrightarrow \phi_p$ è la classica trasformata di Fourier. Così come è scritta è una relazione formale che acquista significato quando si precisano i valori che possono assumere le ϕ_p (ad es $y(x)$ deve essere continua con $\int_0^L y(x)^2 dx < \infty$). Notiamo che l'analogo della matrice Ω^2 è ora costituito dall'operatore differenziale

$$\mathcal{D}^2 = -\left(\frac{\partial}{\partial x}\right)^2$$

Possiamo allora considerare la trasformazione $y(x) \rightarrow \phi_p$ come una rotazione dello spazio ∞ -dimensionale che trasforma l'operatore \mathcal{D}^2 nella matrice diagonale

$$\mathcal{D}^2 \rightarrow (\pi/L)^2 \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 4 & 0 & \dots \\ 0 & 0 & 9 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Questa rotazione lascia invariante il prodotto scalare il che si traduce nella relazione ben nota per la trasformata di Fourier

$$\int_0^L y(x)^2 dx = \sum_{p=1}^{\infty} \phi_p^2$$

e si può attribuire alla radice quadrata di entrambi i membri dell'equazione il significato di *lunghezza* del vettore che rappresenta la forma della fune. Si tratta del primo esempio esplicito di *spazio di Hilbert* che verrà studiato nella Parte III.

1.4.2. Un'equazione integrale. Un altro metodo per studiare le vibrazioni elastiche di una fune porta ad una equazione solo apparentemente diversa dall'equazione d'onda. Supponiamo di applicare una forza \mathbf{f} trasversalmente ad una fune elastica tesa tra due punti fissi. Se si assume che la forza sia sufficientemente debole da causare una

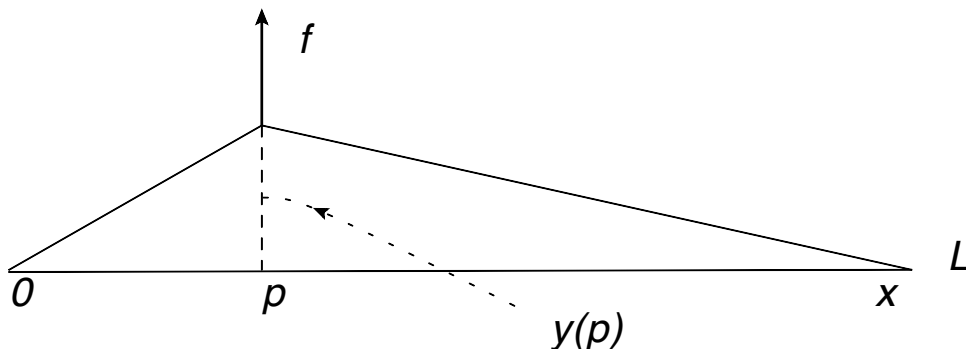


FIGURA 1-4. Deformazione della fune sotto una forza applicata in un punto p .

piccola deformazione, la corda risulta in equilibrio se, detto $y(p)$ lo spostamento della fune nel punto p di applicazione della forza e detta τ la tensione della fune si ha

$$f = \tau y(p)/p + \tau y(p)/(L - p)$$

(si considerino molto piccoli gli angoli in 0 e L). Lo spostamento in un punto qualunque x sarà dato da

$$y(x) = \begin{cases} \frac{x}{p} y(p) & \text{per } x < p \\ \frac{L-x}{L-p} y(p) & \text{per } p < x < L. \end{cases}$$

e in definitiva

$$y(x) = \frac{f}{\tau L} \times \begin{cases} x(L-p) & \text{per } x < p \\ p(L-x) & \text{per } p < x < L. \end{cases}$$

Si noti la forma simmetrica assunta da quest'ultima espressione nello scambio tra x e p . Applicando n forze f_1, f_2, \dots, f_n nei punti p_1, p_2, \dots, p_n e sommando tutti gli spostamenti (cosa lecita trattandosi di piccole deformazioni elastiche) si ottiene

$$y(x) = \frac{1}{\tau L} \left(\sum_{p_k < x} f_k (L - p_k) x + \sum_{p_k > x} f_k (L - x) p_k \right).$$

Supponiamo che n sia molto grande in modo da poter considerare la forza distribuita con continuità sulla fune. Per potere procedere al limite conviene definire $f_k = f(p_k)/n$ e porre $L/n \rightarrow d\xi$, cosicch  si ottiene la semplice formula

$$(1.7) \quad y(x) = \frac{1}{\tau L^2} \int_0^L d\xi f(\xi) \{ (L - \xi) x \vartheta(\xi - x) + (L - x) \xi \vartheta(x - \xi) \}$$

con $\theta(x) = 1$ se $x > 0$ e 0 altrimenti (funzione di Heavyside). Nota la deformazione $y(x)$, per risalire alla forza $f(x)$ distribuita lungo la fune bisogna risolvere un'equazione del tipo

$$y(x) = \int_0^L G(x, \xi) f(\xi) d\xi$$

ovvero un'equazione integrale. Tenendo conto della relazione

$$\frac{d\theta(x)}{dx} = \delta(x)$$

si trova facilmente che

$$\frac{\partial^2 y(x)}{\partial x^2} = -f(x)/(\tau L^2)$$

il che significa che l'operatore differenziale $\partial^2/\partial x^2$ e l'operatore integrale (1.7) sono uno l'inverso dell'altro! La funzione $G(x, \xi)$ prende il nome di *funzione di Green*, termine purtroppo utilizzato in diversi contesti con diversi significati. Se anziché studiare l'equilibrio della fune ne studiassimo il moto, al posto della forza dovremmo introdurre (secondo il principio di D'Alembert) il termine $-\rho \frac{\partial^2 y(x)}{\partial t^2}$, essendo ρ la densità di massa. L'equazione del moto è perciò

$$y(x, t) = -\frac{\rho}{L\tau} \int_0^L \frac{\partial^2 y(\xi, t)}{\partial t^2} \{(L - \xi) x \vartheta(\xi - x) + (L - x) \xi \vartheta(x - \xi)\} d\xi$$

Questa equazione è apparentemente equivalente all'equazione d'onda, in vista della proprietà appena verificata, e tuttavia risulta un po' più generale, in quanto descrive anche soluzioni non differenziabili rispetto a x , come ad es. quella della fune "pizzicata" in un solo punto, che presenta un punto angoloso. Vedremo che in generale l'inversione di una operazione di derivazione porterà ad analoghe espressioni integrali.

Problemi

PROB 1. Calcolare il tensore d'inerzia per un cubo omogeneo di lato ℓ e massa m .

PROB 2. La matrice

$$\mathcal{I} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

può rappresentare il tensore d'inerzia di un corpo rigido? (Cenno: il prodotto $\lambda_1 \lambda_2 \lambda_3$ dei tre momenti d'inerzia principali coincide con il determinante di \mathcal{I}).

PROB 3. Calcolare gli assi principali di un corpo rigido avente il seguente tensore di inerzia

$$\mathcal{I} = \begin{pmatrix} 5/4 & 1/4 & -1/\sqrt{8} \\ 1/4 & 5/4 & -1/\sqrt{8} \\ -1/\sqrt{8} & -1/\sqrt{8} & 3/2 \end{pmatrix}$$

PROB 4. Dimostrare la relazione

$$\mathbf{r} \wedge (\boldsymbol{\omega} \wedge \mathbf{r}) = \boldsymbol{\omega} r^2 - \mathbf{r} (\boldsymbol{\omega} \cdot \mathbf{r})$$

(Cenno: $\mathbf{r} \wedge (\boldsymbol{\omega} \wedge \mathbf{r})$ giace nel piano individuato da $\boldsymbol{\omega}$ e da \mathbf{r} , dunque

$$\mathbf{r} \wedge (\boldsymbol{\omega} \wedge \mathbf{r}) = A\boldsymbol{\omega} + B\mathbf{r}.$$

Si proietta la relazione lungo $\boldsymbol{\omega}$ e \mathbf{r} , e ciò permette di ricavare A e B .)

PROB 5. Considerare la superficie (ellissoide) individuata dall'equazione

$$\sum_{ij} \mathcal{I}_{ij} \omega_i \omega_j = 2T = \text{costante}$$

nello spazio tridimensionale di coordinate $(\omega_1, \omega_2, \omega_3)$.

(a) a) Dimostrare che \mathbf{M} è ortogonale al piano tangente all'ellissoide nel punto $\boldsymbol{\omega}$.

(b) b) Calcolare la distanza dal centro dell'ellissoide al piano tangente all'ellissoide in un punto $\boldsymbol{\omega}$.

PROB 6. Considerare tre masse uguali vincolate a scorrere senza attrito su un profilo circolare e collegate a due a due da molle identiche perfettamente elastiche. Risolvere il problema del moto individuando i modi normali - si noti che in questo caso una frequenza risulta essere uguale a zero.

PROB 7. Calcolare esplicitamente i vettori $\xi_j^{(p)}$ della sez.1.3.

PROB 8. Risolvere il sistema di equazioni differenziali lineari

$$\begin{cases} \dot{x}_1(t) = x_1(t) + \lambda x_2(t) \\ \dot{x}_2(t) = x_2(t) . \end{cases}$$

PROB 9. Risolvere il sistema di equazioni differenziali lineari

$$\begin{cases} \dot{x}_1(t) = \frac{1}{\sqrt{2}} x_2(t) \\ \dot{x}_2(t) = \frac{1}{\sqrt{2}} (x_1(t) + x_3(t)) \\ \dot{x}_3(t) = \frac{1}{\sqrt{2}} x_2(t) . \end{cases}$$

PROB 10. Calcolare la soluzione della equazione d'onda di D'Alembert

$$\frac{\partial^2 \eta(x, t)}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \eta(x, t)}{\partial t^2}$$

sapendo che $\eta(x, 0) = f(x)$, $\left(\frac{\partial \eta}{\partial t}\right)_{t=0} = 0$. (Cenno: sfruttare la soluzione generale $\eta(x, t) = f(x - ct) + g(x + ct)$).

PROB 11. Supponiamo che una corda vibrante sia inizialmente configurata con $\eta(x, 0) = Ax(L-x)$, $\dot{\eta}(x, 0) = 0$. Quali modi normali sono eccitati e con quali ampiezze?

PROB 12. Come nel precedente esercizio, ma

$$\eta(x, 0) = \begin{cases} Ax & \text{per } x < \ell \\ A\ell \frac{L-x}{L-\ell} & \text{per } \ell < x < L \end{cases} .$$

PROB 13. Ad uno scaffale di legno sono appesi N pendoli identici che risultano accoppiati tra loro in quanto lo scaffale può oscillare orizzontalmente nel piano dei pendoli con una forza di richiamo elastica K . Determinare i modi normali del sistema accoppiato scaffale-pendoli nel caso di piccole oscillazioni debolmente smorzate dall'attrito.

Soluzioni

#13 Si veda [BSRW02] per una trattazione dettagliata. Volendo procedere con le proprie forze si può analizzare il problema facilmente nel limite di piccole oscillazioni. Siano θ_i gli angoli e X la coordinata orizzontale dello scaffale. Allora la Lagrangiana è presto formulata come

$$\mathcal{L} = \frac{M}{2} \dot{X}^2 + \frac{m}{2} \sum_i (\ell \dot{\theta}_i + \dot{X})^2 - \frac{M\Omega^2}{2} X^2 - \frac{mg\ell}{2} \sum_i \theta_i^2$$

che danno le equazioni del moto ($\xi = X/\ell$, $\rho = m/M$)

$$\ddot{\xi}(t) = -\Omega^2 \xi(t) + \rho\omega^2 \sum_i \theta_i$$

$$\ddot{\theta}_i(t) = -\omega^2 \theta_i(t) + \Omega^2 \xi(t) - \rho\omega^2 \sum_i \theta_i(t)$$

Se introduciamo la variabile collettiva $\theta_{\parallel} = \frac{1}{\sqrt{N}} \sum_i \theta_i$ e altre $N-1$ variabili $\theta_{\perp i}$ nel piano ortogonale, si vede che le θ_{\perp} si disaccoppiano completamente e costituiscono un sottospazio in cui l'equazione del moto si riduce a

$$\ddot{\theta}_{\perp i} = -\omega^2 \theta_{\perp i}$$

mentre ξ e θ_{\parallel} sono accoppiati:

$$\ddot{\xi} = -\Omega^2 \xi + \rho\omega^2 \sqrt{N} \theta_{\parallel}$$

$$\ddot{\theta}_{\parallel} = -\omega^2 (1 + N\rho) \theta_{\parallel} + \Omega^2 \sqrt{N} \xi$$

che ammette soluzioni periodiche con frequenze

$$\omega_{\pm}^2 = \frac{1}{2} \left(\Omega^2 + \omega^2(1 + N\rho) \pm \sqrt{(\Omega^2 - \omega^2(1 + N\rho))^2 + 4N\rho\omega^2\Omega^2} \right)$$

Per $\omega \ll \Omega$ le due radici diventano approssimativamente

$$\omega_+ \approx \Omega + \frac{N\rho\omega^2}{4\Omega}$$

$$\omega_- \approx \omega - \frac{N\rho\omega^3}{4\Omega^2}$$

e quindi il modo collettivo con tutti i pendoli che oscillano all'unisono è quello con la frequenza minima. Altra faccenda è quella di studiare le oscillazioni in regime non-lineare. C'è sempre la possibilità di utilizzare il calcolo numerico. Il codice *huygens.m* permette di sperimentare il caso generale con un numero qualunque di pendoli e oscillazioni di ampiezza arbitraria in presenza di attrito. Per raggiungere uno stato stazionario stabile, come per gli orologi a cucù presso un tipico orologiaio nelle valli trentine, bisogna anche prevedere un meccanismo di scappamento che rifornisca il sistema dell'energia dissipata. Il tutto presenta la complicazione di un sistema non-lineare, e il fenomeno interessante

che si presenta in certi regimi di parametri è quello scoperto da Huygens nel '600 (v. art. cit.).

II - Operatori Lineari in spazi a dimensione finita

CAPITOLO 2

Spazi Lineari

2.1. Spazi lineari reali

2.1.1. Assiomi. La teoria degli spazi lineari si può formulare assiomaticamente, in modo simile alla Geometria di Euclide, a partire da un certo numero di postulati fondamentali che verranno incorporati nella seguente definizione:*

DEF. 2.1.1. Un insieme \mathfrak{R} di elementi $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$ è detto costituire uno **spazio lineare reale** se

Def.2.1.1

- i) ad ogni coppia di elementi \mathbf{x}, \mathbf{y} in \mathfrak{R} è associato un terzo elemento $\mathbf{z} \in \mathfrak{R}$ detto *somma* di \mathbf{x} con \mathbf{y} e indicato convenzionalmente con $\mathbf{z} = \mathbf{x} + \mathbf{y}$;
- ii) per ogni $\mathbf{x} \in \mathfrak{R}$ ed ogni numero reale \mathbf{a} esiste un elemento di \mathfrak{R} indicato con $\mathbf{a}\mathbf{x}$ e detto *prodotto* di \mathbf{a} per \mathbf{x} .

Per le operazioni di somma e prodotto valgono i seguenti assiomi:

- I.1) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ (commutatività)
- I.2) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$ (associatività)
- I.3) \mathfrak{R} contiene un elemento denominato “zero” indicato con $\mathbf{0}$ tale che $\mathbf{x} + \mathbf{0} = \mathbf{x}$ per ogni \mathbf{x} .
- I.4) Per ogni $\mathbf{x} \in \mathfrak{R}$ esiste un elemento indicato con $-\mathbf{x}$ tale che $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.
- II.1) $1\mathbf{x} = \mathbf{x}, 0\mathbf{x} = \mathbf{0}$
- II.2) $\mathbf{a}(\mathbf{b}\mathbf{x}) = (\mathbf{a}\mathbf{b})\mathbf{x}$
- III.1) $(\mathbf{a} + \mathbf{b})\mathbf{x} = \mathbf{a}\mathbf{x} + \mathbf{b}\mathbf{x}$
- III.2) $\mathbf{a}(\mathbf{x} + \mathbf{y}) = \mathbf{a}\mathbf{x} + \mathbf{a}\mathbf{y}$

Ecco alcuni esempi di spazi lineari reali.

- (a) \mathbb{R}^3 : lo spazio dei vettori ordinari in tre dimensioni con le regole usuali (del parallelogramma) per la somma di vettori e per il prodotto con numeri reali.
- (b) \mathbb{R}^n : per ogni n intero positivo si considera l'insieme delle n -ple ordinate di numeri reali $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$, ..., con

$$\begin{aligned}\mathbf{x} + \mathbf{y} &= (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\ \mathbf{a}\mathbf{x} &= (ax_1, ax_2, \dots, ax_n)\end{aligned}$$

- (c) \mathbb{R}^∞ : lo spazio delle successioni di numeri reali $(x_1, x_2, \dots, x_n, \dots)$ con le operazioni di somma e prodotto analoghe al caso di \mathbb{R}^n .
- (d) Lo spazio $\mathcal{F}(M)$ delle funzioni a valori reali definite su un insieme arbitrario M

$$\mathbf{x}: M \mapsto \mathbb{R}$$

*Ci basiamo sui testi di Gelfand [Gel89], Shilov [Shi77] e Gantmacher [Gan59].

equipaggiato con le operazioni del tutto naturali

$$(\mathbf{x} + \mathbf{y})(\mathbf{m}) = \mathbf{x}(\mathbf{m}) + \mathbf{y}(\mathbf{m})$$

$$(\mathbf{a}\mathbf{x})(\mathbf{m}) = \mathbf{a}\mathbf{x}(\mathbf{m})$$

Si noterà come quest'ultimo esempio comprenda i precedenti come casi particolari ($\mathbf{M} = \{1, 2, 3\}$, $\mathbf{M} = \{1, 2, \dots, \mathbf{n}\}$, $\mathbf{M} = \mathbb{Z}^+$).

- (e) Lo spazio $\mathcal{P}_{\mathbf{n}}$ dei polinomi \mathbf{p} in una variabile reale o complessa \mathbf{x} con grado minore o uguale a \mathbf{n} :

$$\mathbf{p}(\mathbf{x}) = \sum_{j=0}^{\mathbf{n}} \mathbf{p}_j \mathbf{x}^j$$

(rientra in (d) con $\mathbf{M} = \mathbb{R}$ oppure \mathbb{C} e con la restrizione che le funzioni devono essere polinomi).

- (f) Lo spazio $\mathbf{M}(\mathbf{n}, \mathbb{R})$ delle matrici quadrate a \mathbf{n} righe e \mathbf{n} colonne

$$\mathbf{x} = \|\mathbf{x}_{ij}\|_{i,j=1,2,\dots,\mathbf{n}}$$

$$(\mathbf{x} + \mathbf{y})_{ij} = \mathbf{x}_{ij} + \mathbf{y}_{ij}$$

$$(\mathbf{a}\mathbf{x})_{ij} = \mathbf{a}\mathbf{x}_{ij}$$

Def.2.1.2

DEF. 2.1.2. Sia \mathfrak{R} uno spazio lineare. Si dice che \mathbf{k} vettori $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\mathbf{k}}$ di \mathfrak{R} sono *linearmente dipendenti* se esistono \mathbf{k} numeri reali $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{\mathbf{k}}$ **non tutti nulli** tali che

$$\sum_{j=1}^{\mathbf{k}} \mathbf{a}_j \mathbf{x}_j = \mathbf{0}$$

Se \mathbf{k} vettori non sono linearmente dipendenti, essi si dicono *linearmente indipendenti*: ciò significa che in questo caso la relazione $\sum_{j=1}^{\mathbf{k}} \mathbf{a}_j \mathbf{x}_j = \mathbf{0}$ implica $\mathbf{a}_1 = \mathbf{a}_2 = \dots = \mathbf{a}_{\mathbf{k}} = \mathbf{0}$.

Risulta immediato dalla definizione che se \mathbf{k} vettori $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\mathbf{k}}$ sono linearmente dipendenti, allora almeno uno tra di essi, diciamo $\mathbf{x}_{\mathbf{k}}$, è esprimibile come combinazione lineare degli altri. Il concetto di lineare indipendenza permette di definire la *dimensionalità* di uno spazio lineare. Nel caso più intuitivo dello spazio ordinario (tridimensionale) si ha che qualunque vettore è esprimibile come combinazione lineare di tre vettori *non coplanari*. Dunque nello spazio esistono terne di vettori linearmente indipendenti, ma ogni gruppo di quattro o più vettori risultano linearmente dipendenti. Uno spazio lineare si dice avere dimensione finita se esiste un limite superiore al numero di vettori linearmente indipendenti, e cioè:

Def.2.1.3

DEF. 2.1.3. Uno spazio lineare \mathfrak{R} si dice \mathbf{n} -dimensionale se

- i) esso contiene \mathbf{n} vettori linearmente indipendenti
- ii) $\mathbf{n} + 1$ vettori sono sempre linearmente dipendenti.

Uno spazio lineare si dice infinito-dimensionale se esso contiene \mathbf{n} vettori indipendenti per ogni numero naturale \mathbf{n} .

In questa prima parte considereremo solo spazi lineari a \mathbf{n} dimensioni. Se non specificato diversamente, \mathfrak{R} sarà uno spazio lineare \mathbf{n} -dimensionale.

2.1.2. Basi in uno spazio lineare.

Def.2.1.4

DEF. 2.1.4. Ogni insieme di \mathbf{n} vettori $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\mathbf{n}}$ linearmente indipendenti in uno spazio \mathfrak{R} \mathbf{n} -dimensionale si dice costituire *una base* in \mathfrak{R} .

Assegnata una base in \mathfrak{R} , ogni vettore x si può individuare mediante n numeri reali. Infatti per ipotesi x, x_1, x_2, \dots, x_n sono linearmente dipendenti e dunque esiste una scelta di costanti $\{a_i, i = (0, 1, 2, \dots, n)\}$ tali che

$$-a_0 x + a_1 x_1 + a_2 x_2 + \dots + a_n x_n = 0$$

con $a_0 \neq 0$ (altrimenti x_1, x_2, \dots, x_n sarebbero *linearmente dipendenti*, per cui x è esprimibile come combinazione lineare dei vettori della base)

$$x = a_0^{-1}(a_1 x_1 + a_2 x_2 + \dots + a_n x_n)$$

e la corrispondenza $x \Leftrightarrow \{a_i/a_0\}$ è biunivoca. Dunque uno spazio lineare \mathfrak{R} è in corrispondenza biunivoca con lo spazio delle n -ple ordinate di numeri reali. Tale corrispondenza è ovviamente dipendente dalla base scelta. Se $x = c_1 x_1 + c_2 x_2 + \dots + c_n x_n$, i numeri reali $c_i, i = 1, 2, \dots, n$ si diranno le *componenti* di x nella base (x_1, x_2, \dots, x_n) . È immediato verificare che le operazioni di somma e di prodotto per un numero reale si traducono nella somma delle rispettive componenti e nel prodotto delle componenti per lo stesso numero reale.

$$\begin{aligned} x^{(1)} = \sum_{j=1}^n c_j^{(1)} x_j & \quad x^{(1)} + x^{(2)} = \sum_{j=1}^n (c_j^{(1)} + c_j^{(2)}) x_j \\ \implies & \\ x^{(2)} = \sum_{j=1}^n c_j^{(2)} x_j & \quad a x^{(1)} = \sum_{j=1}^n (a c_j^{(1)}) x_j \end{aligned}$$

Ne segue il

TEOREMA 2.1.1. *Tutti gli spazi lineari n -dimensionali sono tra loro isomorfi. Spazi lineari con differente dimensionalità sono tra loro non isomorfi.*

Teor.2.1.1

(per isomorfismo tra spazi lineari si intende l'esistenza di una corrispondenza biunivoca che rispetta le operazioni di somma e prodotto).

DEF. 2.1.5. Un sottoinsieme $\mathfrak{R}' \subset \mathfrak{R}$ è detto *sottospazio* di \mathfrak{R} se esso forma uno spazio lineare secondo le stesse operazioni definite in \mathfrak{R} .

Def.2.1.5

L'insieme vuoto $\{\emptyset\}$ e $\{\mathfrak{R}\}$ soddisfano banalmente alla definizione e si dicono pertanto sottospazi *impropri*. Ogni altro sottospazio verrà detto *proprio*. L'insieme \mathfrak{R}' di tutte le combinazioni lineari di k vettori linearmente indipendenti forma un sottospazio lineare k -dimensionale.

Siano ora (e_1, e_2, \dots, e_n) e $(e'_1, e'_2, \dots, e'_n)$ due basi in \mathfrak{R} . Si avrà allora

$$\begin{aligned} e'_1 &= a_{11} e_1 + a_{21} e_2 + \dots + a_{n1} e_n \\ e'_2 &= a_{12} e_1 + a_{22} e_2 + \dots + a_{n2} e_n \\ &\vdots \\ e'_n &= a_{1n} e_1 + a_{2n} e_2 + \dots + a_{nn} e_n \end{aligned}$$

La matrice $\|a_{ij}\|$ ha determinante non nullo; infatti se fosse $\det(a) = 0$ per il teorema di Cramer esisterebbe una n -pla di numeri reali non tutti nulli c_1, \dots, c_n tali che

$$\sum_{i=1}^n a_{ij} c_j = 0$$

da cui seguirebbe

$$\sum_j c_j e'_j = \sum_i \sum_j c_i a_{ji} e_j = 0$$

ossia $\{e'_j\}$ non costituirebbe una base. La matrice $\mathbf{A} = \|a_{ij}\|$ si dice la *matrice di trasformazione* dalla base $\{e\}$ alla base $\{e'\}$. La conoscenza di \mathbf{A} permette di esprimere le componenti di un vettore rispetto a una base in termini delle componenti rispetto alla seconda base (ove non ci sia pericolo di ambiguità eviteremo di indicare i limiti delle sommatorie)

$$\begin{aligned} x &= \sum_i c_i e_i = \sum_i c'_i e'_i = \sum_{i,j} c'_i a_{ji} e_j \\ &= \sum_j \left(\sum_i a_{ji} c'_i \right) e_j \end{aligned}$$

da cui

$$c_j = \sum_i a_{ji} c'_i$$



In simbolismo matriciale (adottando la convenzione del prodotto “righe per colonne”) si ha

$$(e'_1 \ e'_2 \ \dots \ e'_n) = (e_1 \ e_2 \ \dots \ e_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_n \end{pmatrix}$$

L'espressione di un vettore x in termini delle due basi si può allora ritrovare in questo formalismo come segue

$$\begin{aligned} x &= (e_1 \ e_2 \ \dots \ e_n) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = (e_1 \ e_2 \ \dots \ e_n) \mathbf{A} \begin{pmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_n \end{pmatrix} \\ &= (e'_1 \ e'_2 \ \dots \ e'_n) \begin{pmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_n \end{pmatrix} \end{aligned}$$

 Il formalismo matriciale è del tutto equivalente all'uso esplicito degli indici. Conviene familiarizzarsi con questo sistema che è alla base della manipolazione di matrici e vettori nel linguaggio `matlab`, che verrà utilizzato nel seguito per risolvere problemi di analisi lineare. 

2.2. Spazi lineari complessi

Quanto detto finora per gli spazi lineari reali si può ripetere con poche modifiche per gli *spazi lineari complessi*, definiti come segue:

DEF. 2.2.1. Un insieme \mathfrak{C} di elementi ξ, η, ζ, \dots è detto costituire uno spazio lineare complesso se

Def.2.2.1

- i) identico alla Def. 2.1.1;
- ii) per ogni $\xi \in \mathfrak{C}$ ed ogni numero complesso α esiste un elemento $\alpha\xi \in \mathfrak{C}$;
- i) e ii) devono soddisfare i seguenti assiomi
 - I) identico alla Def. 2.1.1
 - II) $1\xi = \xi; \alpha(\beta\xi) = (\alpha\beta)\xi$;
 - III) $(\alpha + \beta)\xi = \alpha\xi + \beta\xi, \alpha(\xi + \eta) = \alpha\xi + \alpha\eta$.

Per gli spazi lineari complessi si definiscono per estensione dal caso reale i concetti di *lineare indipendenza*, *dimensionalità*, *base*, *sottospazio*, *componenti* e *trasformazioni di base*.

Notiamo che per ogni spazio lineare complesso n -dimensionale \mathfrak{C} può essere riguardato come uno spazio lineare reale $2n$ -dimensionale. Sia infatti (e_1, e_2, \dots, e_n) una base in \mathfrak{C} e sia $\xi = \sum_i \alpha_i e_i$. La corrispondenza

$$\xi \rightarrow \{\Re\{\alpha_1\}, \Re\{\alpha_2\}, \dots, \Re\{\alpha_n\}, \Im\{\alpha_1\}, \Im\{\alpha_2\}, \dots, \Im\{\alpha_n\}\}$$

definisce un isomorfismo $\mathfrak{C} \mapsto \mathbb{R}^{2n}$. Da questo punto in avanti tratteremo sempre il caso generale di spazi lineari complessi.

2.3. Spazi lineari metrici

Le definizioni date finora permettono di sviluppare la geometria “affine”. È noto che per trattare la geometria Euclidea nello spazio tridimensionale ordinario è necessario introdurre il concetto di lunghezza di un vettore, ovvero quello di *prodotto scalare* (o prodotto “interno”) di due vettori qualunque. A tale scopo si introduce la seguente definizione:

DEF. 2.3.1. Uno *spazio lineare unitario* è uno spazio lineare complesso in cui è definito un prodotto interno, cioè una funzione che associa ad ogni coppia di vettori ξ, η un numero complesso, denotato con $\langle \xi, \eta \rangle$ in modo tale che i seguenti assiomi siano soddisfatti:

Def.2.3.1

- i) $\langle \xi, \eta \rangle = \overline{\langle \eta, \xi \rangle}$;
- ii) $\langle \xi, \alpha\eta \rangle = \alpha\langle \xi, \eta \rangle, \quad (\forall \alpha \in \mathbb{C})$;
- iii) $\langle \xi, \eta_1 + \eta_2 \rangle = \langle \xi, \eta_1 \rangle + \langle \xi, \eta_2 \rangle$;
- iv) $\langle \xi, \xi \rangle \geq 0$ con $\langle \xi, \xi \rangle = 0 \Leftrightarrow \xi = 0$.

In analogia con il caso reale, si indica con $\|\xi\|$ e si chiama *norma* (o *lunghezza*) del vettore ξ la quantità $\|\xi\| = \sqrt{\langle \xi, \xi \rangle}$. Due vettori ξ, η si dicono *ortogonali* se $\langle \xi, \eta \rangle = 0$.

Esempi di spazi unitari:

- a) Sia \mathbb{C}^n lo spazio delle n -ple di numeri complessi $\xi = (z_1, z_2, \dots, z_n)$. Se $\eta = (w_1, w_2, \dots, w_n)$, si definisce il prodotto interno come

$$\langle \xi, \eta \rangle = \sum_j \bar{z}_j w_j.$$

Si controlla facilmente che tutti gli assiomi della Def. 2.3.1 sono soddisfatti.

- b) Sia \mathcal{P}_n lo spazio dei polinomi in una variabile reale di grado inferiore a n e con coefficienti complessi. Si definisce

$$\langle p_1, p_2 \rangle = \int_{-1}^1 \overline{p_1(x)} p_2(x) dx$$

- c) Sia \mathcal{Q}_n lo spazio dei polinomi in una variabile complessa di grado inferiore a n . Si definisce

$$\langle q_1, q_2 \rangle = \iint_{\mathbb{R}^2} \frac{\overline{q_1(x+iy)} q_2(x+iy)}{(1+x^2+y^2)^{n+1}} dx dy.$$

Sia (e_1, e_2, \dots, e_n) una base in \mathfrak{E} . Dalle proprietà del prodotto interno segue che la conoscenza della matrice

$$h_{ij} = \langle e_i, e_j \rangle$$

è sufficiente per calcolare il prodotto interno di qualunque coppia di vettori. Infatti se $\xi = \sum z_i e_i$, $\eta = \sum w_i e_i$, si ha

$$\begin{aligned} \langle \xi, \eta \rangle &= \left\langle \sum z_i e_i, \sum w_j e_j \right\rangle \\ &= \sum_{i,j} \overline{z_i} w_j \langle e_i, e_j \rangle \\ &= \sum_{i,j} \overline{z_i} h_{ij} w_j \end{aligned}$$

La matrice $\|h_{ij}\|$ si dice la *matrice metrica* dello spazio unitario \mathfrak{E} e gode delle seguenti proprietà:

- i) $h_{ij} = \overline{h_{ji}}$, ovvero h è *Hermitiana*;
- ii) per ogni n -pla di numeri complessi $\alpha_1, \dots, \alpha_n$ si ha $\sum_{ij} \overline{\alpha_i} h_{ij} \alpha_j \geq 0$ e il segno di uguaglianza vale solo se tutte le α_j sono nulle. Si dice che h è *positiva definita*.

Ogni matrice che soddisfa i) e ii) può essere utilizzata per definire un prodotto interno.

Teor.2.3.1

TEOREMA 2.3.1. *Ogni spazio unitario ammette basi ortonormali ossia tali che la matrice metrica è data dalla matrice unità $h_{ij} = \delta_{ij}$.*

PROVA. Da una qualunque base (e_1, e_2, \dots, e_n) si costruisce una base ortonormale (f_1, f_2, \dots, f_n) attraverso il procedimento di Schmidt: Sia $f_1 = e_1$, $f_2 = \frac{1}{\|e_2\|} \begin{vmatrix} \langle e_1, e_1 \rangle & \langle e_1, e_2 \rangle \\ e_1 & e_2 \end{vmatrix}$, e in generale per $k \leq n$

$$f_k = \frac{1}{\| \begin{vmatrix} \langle e_1, e_1 \rangle & \langle e_1, e_2 \rangle & \dots & \langle e_1, e_k \rangle \\ \langle e_2, e_1 \rangle & \langle e_2, e_2 \rangle & \dots & \langle e_2, e_k \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle e_{k-1}, e_1 \rangle & \langle e_{k-1}, e_2 \rangle & \dots & \langle e_{k-1}, e_k \rangle \end{vmatrix} \|} \begin{vmatrix} \langle e_1, e_1 \rangle & \langle e_1, e_2 \rangle & \dots & \langle e_1, e_k \rangle \\ \langle e_2, e_1 \rangle & \langle e_2, e_2 \rangle & \dots & \langle e_2, e_k \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle e_{k-1}, e_1 \rangle & \langle e_{k-1}, e_2 \rangle & \dots & \langle e_{k-1}, e_k \rangle \end{vmatrix}$$

Si verifica immediatamente che f_k è ortogonale a tutti i vettori e_1, e_2, \dots, e_{k-1} e quindi a tutti gli f_1, f_2, \dots, f_{k-1} . Resta ora solo da “normalizzare” i vettori della nuova base, ossia si moltiplica ogni f_k per $\|f_k\|^{-1}$. ■

In pratica, avendo costruito la matrice metrica h , per ogni k si prende il minore $\|h_{ij}\|_{i,j \leq k}$ e si sostituisce l'ultima riga con $\{e_1, \dots, e_k\}$.

Si noti che in una base ortonormale le componenti di un vettore sono date dal prodotto scalare con i vettori di base, cioè

$$\begin{aligned}\xi &= \sum_j \alpha_j e_j \\ \langle e_k, \xi \rangle &= \sum_j \alpha_j \langle e_k, e_j \rangle = \alpha_k\end{aligned}$$

2.3.1. Cambiamenti di base e matrici unitarie. Siano (e_1, e_2, \dots, e_n) e $(e'_1, e'_2, \dots, e'_n)$ due basi ortonormali; sarà allora, per quanto appena detto

$$e'_i = \sum_j u_{ji} e_j .$$

La matrice $\mathcal{U} = \|u_{ij}\|$ ha come elementi di matrice $u_{ij} = \langle e_i, e'_j \rangle$. Se scambiamo una base con l'altra troviamo che la matrice \mathcal{U} gode di una notevole proprietà:

$$\overline{u_{ij}} = \overline{\langle e_i, e'_j \rangle} = \langle e'_j, e_i \rangle = (\mathcal{U}^{-1})_{ji}$$

cioè la matrice inversa di \mathcal{U} si ottiene prendendo la coniugata complessa della matrice trasposta, detta in breve la *coniugata Hermitiana* di \mathcal{U} , indicata con \mathcal{U}^\dagger . Si noti che, come si controlla immediatamente ricorrendo alla rappresentazione in componenti, la coniugazione Hermitiana soddisfa l'identità

$$\langle \xi, \mathcal{A} \eta \rangle = \langle \mathcal{A}^\dagger \xi, \eta \rangle .$$

per ogni matrice complessa \mathcal{A} .

DEF. 2.3.2. Una matrice \mathcal{U} si dice *unitaria* se soddisfa la condizione

Def.2.3.2

$$\mathcal{U} \mathcal{U}^\dagger = \mathcal{U}^\dagger \mathcal{U} = \mathbb{1}$$

(il simbolo $\mathbb{1}$ indicherà sempre la matrice unità, la matrice con elementi uguali a uno sulla diagonale e zero altrove, $\mathbb{1}_{ij} = \delta_{ij}$).

La definizione rappresenta la naturale estensione al caso di spazi lineari unitari del concetto di matrice *ortogonale* che, per spazi reali, costituisce la generica matrice di trasformazione da una base ortonormale ad un'altra. Per le matrici ortogonali si ha $\mathcal{O}^{-1} = \mathcal{O}'$, indicando con \mathcal{O}' la matrice trasposta (convenzione adottata anche in `matlab`). Segue immediatamente dalla definizione che il determinante di una matrice unitaria è un numero complesso di modulo uno:

$$|\det(\mathcal{U})|^2 = \det(\mathcal{U}) \det(\mathcal{U}^\dagger) = \det(\mathcal{U} \mathcal{U}^\dagger) = 1 .$$

Si osservi che il prodotto di due matrici unitarie è ancora una matrice unitaria. L'insieme delle matrici unitarie di dimensione n forma pertanto un *gruppo*, essendo verificati anche tutti gli altri assiomi (associatività, esistenza dell'elemento neutro e dell'inverso). A titolo di esempio consideriamo il caso più semplice non banale, quello delle matrici unitarie di dimensione 2, denominato gruppo $\mathcal{U}(2)$. Dalla condizione

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^{-1} = \begin{pmatrix} \overline{\alpha} & \overline{\gamma} \\ \overline{\beta} & \overline{\delta} \end{pmatrix}$$

e indicando con $\Delta = e^{i\varphi}$ il determinante della matrice, segue $\delta = \bar{\alpha} e^{i\varphi}$, $\gamma = -\bar{\beta} e^{i\varphi}$, e perciò ogni matrice unitaria 2×2 è della forma

$$\mathcal{U} = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} e^{i\varphi} & \bar{\alpha} e^{i\varphi} \end{pmatrix}$$

con la condizione $|\alpha|^2 + |\beta|^2 = 1$. Il sottoinsieme caratterizzato da $\Delta = 1$ forma un sottogruppo indicato con $SU(2)$, il gruppo delle matrici unitarie 2×2 con determinante uguale a uno. Una conveniente parametrizzazione è la seguente[†]:

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} = \cos(\theta/2) \mathbb{1} + i \sin(\theta/2) \mathbf{n} \cdot \boldsymbol{\sigma}$$

essendo $\mathbf{n} = (n_1, n_2, n_3)$ un qualunque vettore reale di lunghezza uno (un “versore”) e $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ una tripletta di matrici Hermitiane

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

note come *matrici di Pauli* (introdotte in realtà dal matematico inglese Arthur Cayley decenni prima[‡], e già note sotto forma di algebra dei quaternioni da W.R. Hamilton, 1843).

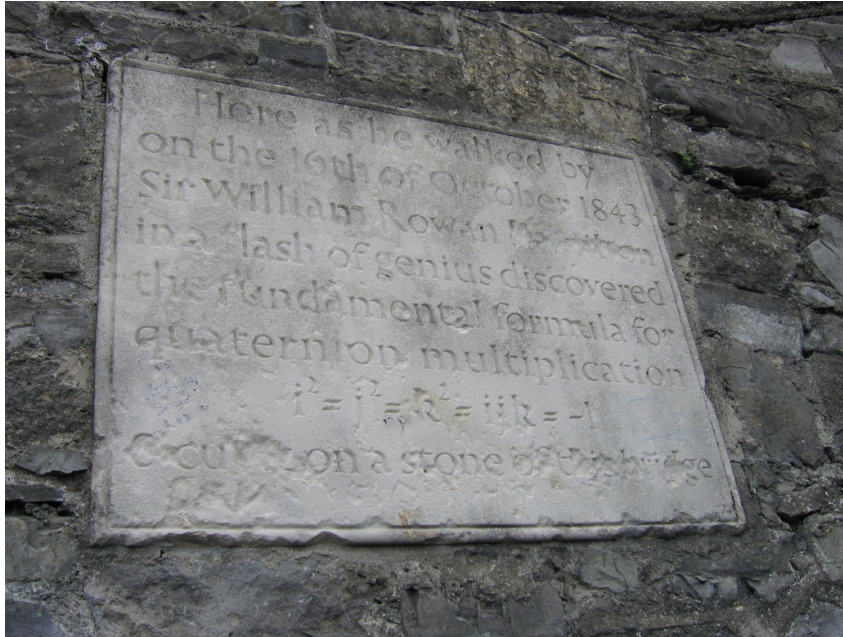


FIGURA 2-1. Brougham Bridge, Dublin, dove i quaternioni sono venuti alla luce.

[†]Equivalente alla rappresentazione in termini di quaternioni, $q \in \mathbb{H}$: $\mathcal{U} \leftrightarrow q_0 + q_1 i + q_2 j + q_3 k$, $\sum q_j^2 = 1$.

[‡]L'opera omnia di Cayley si trova on line attraverso Wikipedia.

PROBLEMA 2-1. *Date due matrici unitarie parametrizzate in termini di (θ, \mathbf{n}) , determinare il loro prodotto, facendo uso della identità [§]*

$$\sigma_j \sigma_k = \delta_{jk} \mathbb{1} + i \sum_m \varepsilon_{jkm} \sigma_m$$

Il gruppo $SU(2)$ è intimamente legato al gruppo delle rotazioni in tre dimensioni reali, ossia al gruppo delle matrici ortogonali $SO(3)$ (vedi Probl. 70).

2.4. Funzionali lineari e spazio duale

Dato uno spazio lineare \mathfrak{C} , consideriamo una applicazione

$$\phi: \mathfrak{C} \mapsto \mathbb{C}$$

cioè una funzione definita in \mathfrak{C} e a valori complessi, che soddisfi alle seguenti proprietà:

$$\begin{aligned} \phi(\xi + \eta) &= \phi(\xi) + \phi(\eta) \\ \phi(\lambda \xi) &= \lambda \phi(\xi) \quad (\lambda \in \mathbb{C}) \end{aligned}$$

ϕ si dice allora un *funzionale lineare* su \mathfrak{C} . Dati due funzionali lineari ϕ_1 e ϕ_2 possiamo considerare la loro combinazione lineare $\phi = \alpha_1 \phi_1 + \alpha_2 \phi_2$ definita nel modo naturale

$$\phi(\xi) = \alpha_1 \phi_1(\xi) + \alpha_2 \phi_2(\xi)$$

Il funzionale nullo $\mathbf{0}: \mathfrak{C} \mapsto \mathbf{0}$ è un funzionale lineare. In definitiva l'insieme di tutti i funzionali lineari su \mathfrak{C} forma a sua volta uno spazio lineare complesso \mathfrak{C}^* che viene chiamato *spazio duale* di \mathfrak{C} .

TEOREMA 2.4.1. *Esiste una corrispondenza biunivoca $\mathfrak{C} \Leftrightarrow \mathfrak{C}^*$; dunque \mathfrak{C}^* è uno spazio lineare della stessa dimensione di \mathfrak{C} .*

Teor.2.4.1

PROVA. Sia (e_1, e_2, \dots, e_n) una base di \mathfrak{C} ; si avrà

$$\phi(\xi) = \phi\left(\sum_i \alpha_i e_i\right) = \sum_i \alpha_i \phi(e_i).$$

Siano $\phi_i (i = 1, \dots, n)$ i funzionali lineari definiti da

$$\phi_i(e_j) = \delta_{ij};$$

Si ha allora

$$\phi(\xi) = \sum_{ij} \alpha_i \phi(e_j) \phi_j(e_i) = \left(\sum_j \phi(e_j) \phi_j \right) (\xi).$$

Ogni funzionale lineare è pertanto esprimibile come combinazione lineare degli n funzionali lineari ϕ_i . ■

La base $\{\phi_i\}$ in \mathfrak{C}^* è detta la base duale rispetto alla base $\{e_i\}$ in \mathfrak{C} . Finora non si è considerata una struttura metrica in \mathfrak{C} . Supponiamo ora che \mathfrak{C} sia uno spazio unitario. Allora si ha

TEOREMA 2.4.2. *Sia \mathfrak{C} uno spazio unitario. Allora in corrispondenza ad ogni $\phi \in \mathfrak{C}^*$ esiste un unico vettore $\phi^\sharp \in \mathfrak{C}$ tale che*

Teor.2.4.2

$$\phi(\xi) = \langle \phi^\sharp, \xi \rangle, \quad (\forall \xi \in \mathfrak{C}).$$

[§] ε è il simbolo di Ricci, definito da $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$, $\varepsilon_{132} = \varepsilon_{213} = \varepsilon_{321} = -1$, $\varepsilon_{ijk} = 0$ negli altri casi (due indici uguali).

PROVA. La dimostrazione è molto semplice se introduciamo una base e poi dimostriamo che la corrispondenza non dipende dalla base. Diamo qui una dimostrazione diversa, che ammette una estensione immediata al caso ∞ -dimensionale. L'insieme dei vettori ξ tali che $\phi(\xi) = 0$ (il *nucleo* di ϕ), forma un sottospazio di \mathfrak{E} , indicato con $\mathcal{N}(\phi)$. Se la dimensione di \mathcal{N} è n allora ϕ è il funzionale nullo e la corrispondenza è banale. Supponiamo dunque $\dim(\mathcal{N}) < n$. Allora necessariamente $\dim(\mathcal{N}) = n - 1$, come si può facilmente dimostrare. Sia $\mathcal{N}^\perp(\phi)$ l'insieme dei vettori ortogonali ad ogni vettore in $\mathcal{N}(\phi)$. Scegliamo un vettore $\eta \in \mathcal{N}^\perp(\phi)$. Vale l'identità

$$(2.1) \quad \phi(\xi) = \frac{\phi(\eta)}{\|\eta\|^2} \langle \eta, \xi \rangle$$

ossia la corrispondenza cercata è

$$\phi^\sharp = \frac{\overline{\phi(\eta)}}{\|\eta\|^2} \eta.$$

Infatti, decomponiamo ξ secondo $\mathcal{N}(\phi)$ e \mathcal{N}^\perp :

$$\xi = \xi_0 + \alpha \eta, \quad \xi_0 \in \mathcal{N}(\phi).$$

La costante α si determina calcolando su ambo i membri il funzionale lineare, $\phi(\xi) = \alpha \phi(\eta)$ e prendendo il prodotto scalare con η : $\langle \eta, \xi \rangle = \alpha \langle \eta, \eta \rangle$. Dividendo membro a membro si elimina α e si ottiene la relazione cercata Eq. (2.1). ■

Notiamo che la corrispondenza $\phi \rightarrow \phi^\sharp$ gode delle proprietà seguenti:

$$\begin{aligned} \phi_1 + \phi_2 &\Leftrightarrow \phi_1^\sharp + \phi_2^\sharp \\ \alpha \phi &\Leftrightarrow \bar{\alpha} \phi^\sharp. \end{aligned}$$

Tale corrispondenza si dice *antilineare*.

2.4.1. La notazione di Dirac. Nel suo libro “The principles of Quantum Mechanics” [Dir29], Dirac ha introdotto un formalismo per indicare funzionali lineari e vettori che è ormai universalmente adottato nelle applicazioni alla Meccanica Quantistica, in quanto offre una notevole facilità di uso e si presta ad automatizzare certe operazioni sui vettori. Conveniamo di indicare i vettori in \mathfrak{E} con un simbolo (“ket”)

$$|\xi\rangle$$

La lettera ξ , o un arbitrario insieme di simboli, posto all'interno del simbolo di ket ha la funzione di identificare il vettore.

⚡ Questa idea, apparentemente di scarso peso, risulta in realtà preziosa nelle applicazioni alla Fisica atomica o nucleare, dove non di rado si trovano espressioni quali

$$\left\langle JM \left| j_1 j_2 m_1 m_2 \right. \right\rangle$$

o anche

$$\left| \begin{array}{c} j_1 j_2 j_3 \\ m_1 m_2 m_3 \\ \mathbf{E} \mathbf{P}_\pm \end{array} \right\rangle$$

che porrebbero altrimenti qualche problema tipografico. ◀

Denotiamo poi un generico funzionale lineare con il simbolo (“bra”)

$$\langle \phi |$$

con la convenzione che il valore che il funzionale assume su un vettore si indica affiancando semplicemente i due simboli

$$\phi(\xi) \equiv \langle \phi | \xi \rangle .$$

Dato che ad ogni funzionale $\phi \in \mathfrak{C}^*$ corrisponde un unico vettore $\phi^\# \in \mathfrak{C}$, per il Teor.2.4.2, il simbolo $\langle \phi | \xi \rangle$ assume lo stesso valore di $(\phi^\#, \xi)$ e dunque possiamo abbandonare completamente la notazione $\langle \cdot, \cdot \rangle$ per il prodotto interno e mantenere l'unica notazione $\langle \cdot | \cdot \rangle$. Vale la proprietà

$$\langle \phi | \xi \rangle = \overline{\langle \xi | \phi \rangle} .$$

In realtà nel caso presente degli spazi finito-dimensionali la considerazione dello spazio duale non sarebbe strettamente necessaria, ma è utile familiarizzarsi adesso con questi concetti, in quanto la loro introduzione sarà indispensabile nel caso ∞ -dimensionale.

La notazione di Dirac presenta notevoli vantaggi formali (e anche pratici). Per fare un esempio, ricordiamo che per ogni vettore vale la decomposizione

$$\xi = \sum_i e_i (e_i, \xi)$$

essendo $\{e_i\}$ i vettori di una base ortonormale. Nella nuova notazione si ha

$$|\xi\rangle = \sum_i |e_i\rangle \langle e_i | \xi \rangle$$

ovvero

$$|\xi\rangle = \left(\sum_i |e_i\rangle \langle e_i| \right) |\xi\rangle$$

che si può interpretare come il fatto che l'espressione formale $\sum_i |e_i\rangle \langle e_i|$, quando applicata a un qualunque vettore, lo lascia inalterato. D'altra parte l'espressione $|\eta\rangle \langle \eta|$ ha il significato di *proiettore* lungo la direzione individuata dal vettore $|\eta\rangle$, infatti secondo la convenzione di Dirac, l'espressione $|\eta\rangle \langle \eta | \xi \rangle$ rappresenta il vettore $|\eta\rangle$ moltiplicato per il numero complesso $\langle \eta | \xi \rangle$. In conclusione *la somma dei proiettori lungo i vettori di una base ortonormale lascia invariato ogni vettore dello spazio*. Questo fatto si dice costituire la “relazione di completezza” per i vettori $|e_i\rangle$. Vedremo meglio tutto ciò più avanti in termini di “operatori lineari”. Tenendo presente che, nel senso specificato prima, $\sum_i |e_i\rangle \langle e_i| = \mathbb{1}$, si ha

$$\begin{aligned} \langle \eta | \xi \rangle &= \langle \eta | \sum_i |e_i\rangle \langle e_i| \xi \rangle \\ &= \sum_i \overline{\langle e_i | \eta \rangle} \langle e_i | \xi \rangle \end{aligned}$$

che riproduce il prodotto interno in termini di componenti.

Sia ora $\{e'_i\}$ una seconda base ortonormale in \mathfrak{C} , allora

$$\begin{aligned} |e'_i\rangle &= \sum_j |e_j\rangle \langle e_j | e'_i \rangle \equiv \sum_j |e_j\rangle a_{ji} \\ |e_i\rangle &= \sum_j |e'_j\rangle \langle e'_j | e_i \rangle \equiv \sum_j |e'_j\rangle a_{ji}^{-1} \end{aligned}$$

dunque $\mathbf{a}_{ji} = \langle \mathbf{e}_j | \mathbf{e}'_i \rangle$ è la matrice di trasformazione dalla base $\{\mathbf{e}_i\}$ alla $\{\mathbf{e}'_i\}$. Se $\{\mathbf{e}''_i\}$ è una terza base, si avrà

$$\mathbf{a}_{ij} = \langle \mathbf{e}_i | \mathbf{e}'_j \rangle = \sum_k \langle \mathbf{e}_i | \mathbf{e}''_k \rangle \langle \mathbf{e}''_k | \mathbf{e}'_j \rangle = (\mathcal{A}^{e \rightarrow e''} \mathcal{A}^{e'' \rightarrow e})_{ij}$$

ovvero le matrici di trasformazione da una base all'altra si compongono secondo il prodotto “righe-per-colonne”. Indubbiamente il formalismo di Dirac permette di ricavare queste relazioni senza sforzo.

CAPITOLO 3

Operatori Lineari

3.1. Definizioni fondamentali

Abbiamo studiato e classificato nel capitolo precedente (2.4) le applicazioni lineari $\phi : \mathfrak{C} \mapsto \mathbb{C}$. È possibile estendere lo studio di tali funzioni al caso di corrispondenze più generali del tipo $\phi : \mathfrak{C} \mapsto \mathfrak{C}'$ dove \mathfrak{C} e \mathfrak{C}' sono spazi lineari complessi di qualunque dimensione n e n' . Tuttavia ci limiteremo al caso speciale in cui \mathfrak{C}' coincida con \mathfrak{C} (i Matematici parlano allora di *endomorfismi* dello spazio lineare \mathfrak{C}).

DEF. 3.1.1. Una applicazione $\mathcal{A} : \mathfrak{C} \mapsto \mathfrak{C}$ che associa ad ogni vettore $|\xi\rangle \in \mathfrak{C}$ un vettore $\mathcal{A}|\xi\rangle \in \mathfrak{C}$ è detta un *operatore lineare* in \mathfrak{C} se le seguenti condizioni sono soddisfatte:

Def.3.1.1

- i) $\mathcal{A}(|\xi\rangle + |\eta\rangle) = \mathcal{A}|\xi\rangle + \mathcal{A}|\eta\rangle$
- ii) $\mathcal{A}(\lambda|\xi\rangle) = \lambda\mathcal{A}|\xi\rangle, (\forall \lambda \in \mathbb{C})$

Consideriamo alcuni esempi di operatori lineari.

- a) Sia \mathbb{R}^3 lo spazio tridimensionale reale; $\mathcal{R}(\hat{n}, \alpha)$ sia la rotazione di un angolo α intorno al versore \hat{n} . Allora la rotazione è realizzata sui vettori dalla trasformazione

$$(3.2) \quad \mathcal{R}(\mathbf{n}, \alpha) \mathbf{x} = \cos \alpha \mathbf{x} + \sin \alpha \mathbf{n} \wedge \mathbf{x} + (1 - \cos \alpha)(\mathbf{n} \cdot \mathbf{x}) \mathbf{n}.$$

\mathcal{R} è un operatore lineare in \mathbb{R}^3 .

- b) Sia $|\eta\rangle$ un vettore di lunghezza uno in \mathfrak{C} . L'espressione

$$\mathbb{P} = |\eta\rangle \langle \eta|$$

definisce \mathbb{P} come operatore lineare intendendo per convenzione che per ogni $|\xi\rangle \in \mathfrak{C}$ si abbia

$$\mathbb{P}|\xi\rangle = |\eta\rangle \langle \eta|\xi\rangle$$

\mathbb{P} è chiamato *il proiettore* lungo $|\eta\rangle$ in quanto ogni vettore viene trasformato in un multiplo di $|\eta\rangle$. È lo stesso concetto di proiezione di un vettore lungo una direzione che si ha in Geometria Euclidea.

- c) Sia \mathbb{C}^n lo spazio delle n -ple complesse e sia $\|\mathbf{a}_{ij}\|$ una qualunque matrice $n \times n$. Allora la trasformazione

$$(\mathcal{A}\xi)_i = \sum_j \mathbf{a}_{ij} \xi_j$$

definisce \mathcal{A} come un operatore lineare.

- d) Sia \mathcal{P}_n lo spazio dei polinomi nella variabile complessa z aventi grado inferiore o uguale a n . Sia

$$\mathcal{D}p(z) = \frac{dp(z)}{dz}$$

\mathcal{D} è un operatore lineare.

e) L'operatore *identità* $\mathbb{1}$ e l'operatore nullo $\mathbb{0}$, definiti da

$$\mathbb{1} |\xi\rangle = |\xi\rangle ; \mathbb{0} |\xi\rangle = \mathbf{0}.$$

sono i casi estremi di operatori lineari.

⚡ Esistono molti “zeri” da tenere concettualmente distinti: il numero zero 0 , il vettore nullo $\mathbf{0}$, l'operatore nullo $\mathbb{0}$, etc. Ove non si creino ambiguità utilizzeremo in tutti i casi, per non appesantire troppo il formalismo, il simbolo 0 . ◀

Teor.3.1.1

TEOREMA 3.1.1. Ogni operatore lineare è individuato dalla sua azione sui vettori di una base.

PROVA. Sia \mathcal{A} un operatore lineare, e sia $|e_i\rangle \mid i = 1, \dots, n\}$ una base. Sia inoltre

$$\mathcal{A} |e_j\rangle = \sum_i a_{ij} |e_i\rangle$$

Allora si ha

$$\mathcal{A} |\xi\rangle = \mathcal{A} \sum_j \xi_j |e_j\rangle = \sum_{ij} a_{ij} \xi_j |e_i\rangle$$

L'azione di \mathcal{A} è perciò “rappresentata” dalla matrice a_{ij} sulle componenti dei vettori e diremo che $\|a_{ij}\|$ è la *matrice rappresentativa* di \mathcal{A} nella base prescelta. ■

Sia \mathfrak{C} uno spazio unitario. Allora le coordinate di $\mathcal{A}|i\rangle$ nella base ortonormale $|i\rangle$ sono date da

$$\langle j | \mathcal{A} | \xi \rangle = \sum_i \langle j | \mathcal{A} | i \rangle \langle i | \xi \rangle.$$

Dunque in una base ortonormale la matrice rappresentativa di qualunque operatore lineare \mathcal{A} è data da $A_{ij} = \langle i | \mathcal{A} | j \rangle$.

3.1.1. Cambiamenti di base. Sotto un cambiamento di base ortonormale $|\alpha\rangle = \sum_j |j\rangle \langle j | \alpha \rangle = |j\rangle S_{j\alpha}$, la matrice rappresentativa di ogni operatore viene modificata attraverso una *trasformazione di similitudine*:

$$A'_{\alpha\beta} = \langle \alpha | \mathcal{A} | \beta \rangle = \sum_{ij} \langle \alpha | i \rangle \langle i | \mathcal{A} | j \rangle \langle j | \beta \rangle = \sum_{ij} U_{\alpha i} A_{ij} \overline{U_{\beta j}}$$

La matrice rappresentativa nella base “greca” è data quindi da

$$(3.3) \quad A' = U A U^\dagger \equiv U A U^{-1}$$

L'espressione $U A U^{-1}$ è più generale in quanto vale anche nel caso di basi non-ortonormali. Due matrici legate tra loro dalla relazione (3.3) si dicono *matrici simili* e condividono numerose proprietà, come vedremo in seguito.

ESEMPIO 3-1. Consideriamo l'operatore \mathcal{D} introdotto in precedenza (esempio d). Il prodotto interno è definito in modo che

$$|k\rangle \leftarrow \binom{n}{k}^{\frac{1}{2}} z^k, \quad (i = 0, 1, \dots, n)$$

sia una base ortonormale. La matrice rappresentativa di \mathcal{D} si calcola facilmente:

$$D |k\rangle \leftarrow \binom{n}{k}^{\frac{1}{2}} k z^{k-1} \hookrightarrow \binom{n}{k}^{1/2} \binom{n}{k-1}^{-1/2} k |k-1\rangle = \sqrt{k(n-k+1)} |k-1\rangle$$

da cui

$$\langle h | \mathcal{D} | k \rangle = \sqrt{k(n-k+1)} \delta_{h,k-1}$$

Se poniamo $n = 2j$, $k = m+j$, troviamo che gli elementi di matrice divengono $\sqrt{(j+m)(j-m+1)}$ che coincidono con quelli della componente $J_x - iJ_y$ del momento angolare in meccanica quantistica [LL76]. ◀

3.1.2. Algebra degli operatori lineari. Se consideriamo l'insieme $\mathfrak{B}(\mathfrak{C})$ di tutti gli operatori lineari in uno spazio n -dimensionale \mathfrak{C} , troviamo che è possibile definire su tale insieme una struttura algebrica. Innanzitutto \mathfrak{B} è uno spazio lineare n^2 -dimensionale; ciò si riconosce dal fatto che in modo del tutto naturale si possono definire le operazioni di somma e di moltiplicazione per numeri complessi:

$$(\mathcal{A} + \mathcal{B}) |\xi\rangle \equiv \mathcal{A} |\xi\rangle + \mathcal{B} |\xi\rangle$$

$$(\lambda \mathcal{A}) |\xi\rangle \equiv \lambda \mathcal{A} |\xi\rangle$$

La matrice rappresentativa della somma $\mathcal{A} + \mathcal{B}$ in una qualunque base è data dalla somma delle rispettive matrici. Ora, ogni operatore lineare si può esprimere come combinazione lineare degli n^2 operatori $\mathbb{E}^{(ij)}$ aventi rappresentazione matriciale

$$\langle k | \mathbb{E}^{(ij)} | h \rangle \equiv \delta_{ki} \delta_{jh}$$

cioè la matrice $\mathbb{E}^{(ij)}$ ha tutti elementi nulli, tranne quello nella k -esima riga, h -esima colonna. Ovviamente

$$A_{ij} = \sum_{kh} A_{kh} \delta_{ki} \delta_{hj} = \sum_{kh} A_{kh} E_{ij}^{(kh)}$$

e dunque $\mathcal{A} = \sum A_{kh} \mathbb{E}^{(kh)}$.

Si può inoltre definire il prodotto di due operatori come segue

$$(\mathcal{A}\mathcal{B}) |\xi\rangle \equiv \mathcal{A}(\mathcal{B} |\xi\rangle)$$

⚡ L'ordine dei fattori nel prodotto di operatori è essenziale, in generale infatti $\mathcal{B}\mathcal{A} \neq \mathcal{A}\mathcal{B}$, così come è vero per la moltiplicazione delle matrici.

⚡ Gli operatori di rotazione (esempio a a Pag. 35) offrono un esempio importante di operatori non-commutanti. Per esercizio controllare ad es. che rotazioni intorno all'asse x e rotazioni intorno all'asse y di un angolo retto non commutano tra loro.

⚡ Il prodotto tra operatori lineari si traduce in prodotto righe \times colonne delle rispettive matrici rappresentative. Infatti

$$\langle i | \mathcal{A}\mathcal{B} | j \rangle = \sum_k \langle i | \mathcal{A} | k \rangle \langle k | \mathcal{B} | j \rangle$$

ossia $(\mathcal{A}\mathcal{B})_{ij} = \sum_k A_{ik} B_{kj}$. ◀

3.1.3. Funzioni di operatori lineari. Le operazioni definite in \mathfrak{B} permettono di definire *polinomi di un operatore*: se $p(x) = \sum a_k x^k$ si definisce

$$p(\mathcal{A}) \equiv a_0 \mathbb{1} + a_1 \mathcal{A} + a_2 \mathcal{A}^2 + \dots + a_n \mathcal{A}^n$$

essendo per convenzione $\mathcal{A}^0 = \mathbb{1}$, $\mathcal{A}^1 = \mathcal{A}$, ..., $\mathcal{A}^{k+1} = \mathcal{A} \mathcal{A}^k$. L'obiettivo, interessante per le applicazioni, è quello di dare significato a $f(\mathcal{A})$ quando $f(z)$ appartenga ad una determinata classe di funzioni, ad es. *f analitica*. La prima estensione è quella alle funzioni analitiche intere (cioè prive di singolarità) per le quali si utilizza la serie di Taylor

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \Rightarrow f(\mathcal{A}) \equiv \sum_{n=0}^{\infty} a_n \mathcal{A}^n.$$

Ciò è di immediato interesse per le applicazioni, ad es. il sistema differenziale

$$\frac{d}{dt} |\xi, t\rangle = \mathcal{A} |\xi, t\rangle$$

ovvero in componenti

$$\frac{d}{dt} \begin{pmatrix} \xi_1(t) \\ \xi_2(t) \\ \xi_3(t) \\ \vdots \\ \xi_n(t) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & & & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \xi_1(t) \\ \xi_2(t) \\ \xi_3(t) \\ \dots \\ \xi_n(t) \end{pmatrix}$$

ha come soluzione formale

$$|\xi, t\rangle = \exp\{t\mathcal{A}\} |\xi, 0\rangle$$

dove $\exp\{t\mathcal{A}\} \equiv \sum_0^\infty t^n \mathcal{A}^n / n!$. Si pone il problema di controllare la convergenza di una serie di potenze in un operatore lineare. A tale scopo conviene introdurre il concetto di *norma*:

Def.3.1.2

DEF. 3.1.2. Si dice che l'applicazione $\mathcal{A} \in \mathfrak{B} \mapsto \|\mathcal{A}\| \in \mathbb{R}^+$ definisce una norma se sono soddisfatte le seguenti condizioni

- i) $\|\mathcal{A}\| \geq 0$, ($\|\mathcal{A}\| = 0 \Rightarrow \mathcal{A} = \mathbb{O}$)
- ii) $\|\lambda\mathcal{A}\| = |\lambda| \|\mathcal{A}\|$ ($\lambda \in \mathbb{C}$)
- iii) $\|\mathcal{A} + \mathcal{B}\| \leq \|\mathcal{A}\| + \|\mathcal{B}\|$
- iv) $\|\mathcal{A}\mathcal{B}\| \leq \|\mathcal{A}\| \|\mathcal{B}\|$

Una possibile definizione di norma è data da

$$\|\mathcal{A}\| = \sup \frac{\|\mathcal{A}\xi\|}{\|\xi\|} = \max_{\langle \xi | \xi \rangle = 1} \|\mathcal{A}|\xi\rangle\|$$

⚠ L'uso del simbolo $\|\cdot\|$ in vari contesti potrebbe ingenerare confusione: usiamo $\|a_{ij}\|$ per indicare sinteticamente una matrice A , $\|\xi\|$ per la norma del vettore ξ e ora per indicare la norma di un operatore. Dal contesto dovrebbe essere chiaro a quale caso di faccia riferimento, comunque si richiede al lettore una certa attenzione. Il vantaggio è quello di non introdurre un numero eccessivo di notazioni diverse. ◀

Teor.3.1.2

TEOREMA 3.1.2. Una serie di potenze $\sum_0^\infty a_n z^n$ con raggio di convergenza ρ dà luogo a una serie convergente $\sum_0^\infty a_n \mathcal{A}^n$ secondo la norma $\|\cdot\|$ se $\|\mathcal{A}\| < \rho$.

PROVA. Dalle proprietà della norma (iii – iv) segue che

$$\left\| \sum_N^M a_k \mathcal{A}^k \right\| \leq \sum_N^M |a_k| \|\mathcal{A}^k\| \leq \sum_N^M |a_k| \|\mathcal{A}\|^k$$

e dunque è soddisfatta la condizione di Cauchy. ■

Nel caso di funzioni “intere”, cioè con la sola singolarità all'infinito, la serie di Taylor converge per ogni operatore lineare. Vedremo in seguito un metodo generale per il calcolo di funzioni di operatore basato sulla “teoria spettrale”. Per il momento notiamo che, essendo $\mathfrak{B}(\mathfrak{E})$ uno spazio a dimensione finita, nella successione $\{\mathcal{A}^j | j = 0, 1, 2, \dots, k, \dots\}$ da un certo k in poi le potenze di \mathcal{A} saranno linearmente dipendenti dalle potenze più basse. Ciò permette in alcuni casi di sommare rapidamente la serie di Taylor. Si mostrerà più avanti che esiste per ogni operatore lineare \mathcal{A} un polinomio $P(z)$ di ordine non superiore a

$n = \dim(\mathfrak{C})$ tale che $P(\mathcal{A}) = 0$; questo polinomio si calcola facilmente una volta introdotta una rappresentazione in una base e cioè $P(z) = \det(z\mathbb{1} - A)$ e risulta indipendente dalla base prescelta.

PROBLEMA 3-2. Calcolare la funzione $\exp\{i t \sigma_1\}$ essendo σ_1 la prima matrice di Pauli (vedi il § 2.3.1 a Pag. 30).

SOLUZIONE. La serie $\sum_0^\infty (it)^n \sigma_1^n / n!$ si può facilmente risommare osservando che $\sigma_1^2 = \mathbb{1}$ e quindi

$$\begin{aligned} \sum_0^\infty \frac{(it)^n}{n!} \sigma_1^n &= \sum_{n=0}^\infty \frac{(-)^n t^{2n}}{(2n)!} \mathbb{1} + i \sum_{n=0}^\infty \frac{(-)^n t^{2n+1}}{(2n+1)!} \sigma_1 \\ &= \cos t \mathbb{1} + i \sin t \sigma_1 \\ &= \begin{pmatrix} \cos t & i \sin t \\ i \sin t & \cos t \end{pmatrix} \end{aligned}$$

◀

PROBLEMA 3-3. Lo stesso problema precedente ma per la matrice

$$J = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Suggerimento: cercare una relazione che legghi J^3 alle potenze inferiori. Provare a utilizzare il risultato anticipato nel testo, intermini di determinante.

PROBLEMA 3-4. Si calcoli la funzione $f(\sigma_1)$ per

- (a) $f(z) = \sin(z)$
- (b) $f(z) = 1/(1 + \varepsilon z)$
- (c) $f(z) = \log(1 + \varepsilon z)$

🔗 Si sarà notato, nell'ultimo problema, che si sono considerate anche funzioni non intere, e quindi la convergenza non vale incondizionatamente (vincolo su ε). La funzione $(\mathbb{1} + A)^{-1}$ ritorna, se esiste e se la serie converge, giusto l'operatore inverso di $\mathbb{1} + A$. Questo caso particolare rientra in uno ben più generale che vedremo in seguito. ▶

🔗 Il linguaggio **matlab** offre un ausilio molto potente per manipolare matrici e vettori, costituendo un vero e proprio laboratorio di calcolo. Assumiamo che il lettore sia familiare con le basi del linguaggio e le convenzioni adottate. In **matlab** la funzione esponenziale di matrici è realizzata dalla funzione **expm**, ad es.

```
>> s1=[0,1;1,0];
>> expm(i*pi/2*s1)
ans =
      0      0 + 1.0000i
0 + 1.0000i      0
```

Tuttavia può risultare più efficiente risolvere direttamente l'equazione

$$\frac{d}{dt} |t\rangle = i \sigma_1 |t\rangle$$

attraverso la routine **ode45**: si definisce una funzione

```
function y=pauli(t,x)
s1=[0,1;1,0];
y=i*s1*x;
```

e si trova in un colpo solo il valore di $\exp\{it\sigma_1\}$ con il comando

```
>> t=linspace(0,2*pi,128);
>> [T,firstcolumn]=ode45(@pauli,t,[1;0]);
>> [T,secondcolumn]=ode45(@pauli,t,[0;1]);
```

ode45 è solo una delle numerose routines di soluzione numerica di equazioni differenziali ordinarie; si consultì la documentazione per scegliere quella più adatta al proprio problema. Caso per caso sarà chiaro se interessa l'intera matrice $\exp\{\mathcal{A}\}$ oppure la sua azione su un singolo vettore e si sceglierà di conseguenza la routine più economica. Se si desidera un risultato esatto si può utilizzare un linguaggio simbolico quale *Mathematica*:

```
In[1]:= MatrixExp[I t {{0, 1}, {1, 0}}] // MatrixForm

Out[1]//MatrixForm=

$$\begin{pmatrix} \cos[t] & i \sin[t] \\ i \sin[t] & \cos[t] \end{pmatrix}$$

```

FIGURA 3-1. Calcolo di $\exp m$ con *Mathematica* (1).

La vita è comoda: anche il secondo esercizio si risolve senza fatica

```
In[2]:= MatrixExp[I t {{0, 1, 0}, {1, 0, 1}, {0, 1, 0}} / Sqrt[2]] // MatrixForm

Out[2]//MatrixForm=

$$\begin{pmatrix} \frac{1}{2} (1 + \cos[t]) & \frac{i \sin[t]}{\sqrt{2}} & \frac{1}{2} (-1 + \cos[t]) \\ \frac{i \sin[t]}{\sqrt{2}} & \cos[t] & \frac{i \sin[t]}{\sqrt{2}} \\ \frac{1}{2} (-1 + \cos[t]) & \frac{i \sin[t]}{\sqrt{2}} & \frac{1}{2} (1 + \cos[t]) \end{pmatrix}$$

```

FIGURA 3-2. Calcolo di $\exp m$ con *Mathematica* (2).

Tuttavia conviene agli inizi allenarsi a calcolare senza questi ausili automatici – inoltre non è consigliabile affidarsi acriticamente a questi strumenti! ◀

3.1.4. L'operatore inverso.

DEF. 3.1.3. Un operatore lineare \mathcal{A} si dice *invertibile* se esiste un operatore lineare denotato con \mathcal{A}^{-1} tale che per ogni $|\rangle \in \mathfrak{C}$ si abbia (per economia, indicheremo con il ket senza etichetta un generico vettore dello spazio \mathfrak{C})

Def.3.1.3

$$\mathcal{A} \mathcal{A}^{-1} |\rangle = \mathcal{A}^{-1} \mathcal{A} |\rangle = |\rangle .$$

\mathcal{A}^{-1} , se esiste, si dirà *l'inverso* di \mathcal{A} .

TEOREMA 3.1.3. Un operatore $\mathcal{A} \in \mathfrak{B}(\mathfrak{C})$ è invertibile se e solo se esso definisce una trasformazione biunivoca di \mathfrak{C} in sè.

Teor.3.1.3

PROVA. Supponiamo che $\mathcal{A}|1\rangle = \mathcal{A}|2\rangle$; l'esistenza dell'inverso implica $\mathcal{A}^{-1} \mathcal{A}|1\rangle = |1\rangle = \mathcal{A}^{-1} \mathcal{A}|2\rangle = |2\rangle$. Viceversa se da $\mathcal{A}|1\rangle = \mathcal{A}|2\rangle$ discende $|1\rangle = |2\rangle$ si può dedurre che $\mathcal{A}|\star\rangle = \mathbf{0}$ non ammette soluzioni tranne il vettore nullo. Sia allora $|f_i\rangle = \mathcal{A}|e_i\rangle$ per una base di vettori $\{|e_i\rangle\}$. Allora anche $\{|f_i\rangle\}$ costituisce una base, in quanto $\mathbf{0} = \sum_i \alpha_i |f_i\rangle = \mathcal{A}(\sum_i \alpha_i |e_i\rangle)$ impone che tutti gli α siano nulli. Si definisce allora

$$\mathcal{A}^{-1} |f_i\rangle = |e_i\rangle$$

e si estende \mathcal{A}^{-1} per linearità a tutto lo spazio. ■

La matrice rappresentativa della matrice \mathcal{A}^{-1} , se esiste, è semplicemente la matrice inversa della matrice che rappresenta \mathcal{A} . Ricordiamo come si caratterizza una matrice invertibile e come si costruisce la sua inversa. Assegnata una matrice $\mathbf{A} = \|\mathbf{a}_{ij}\|$ vale la seguente proprietà (sviluppo di Lagrange)

$$\sum_k \mathbf{a}_{ik} \mathfrak{A}_{jk} = \det(\mathbf{A}) \delta_{ij}$$

essendo \mathfrak{A}_{jk} il *complemento algebrico* corrispondente all'elemento \mathbf{a}_{jk} . Questa relazione vale qualunque sia il valore del determinante; se $\det(\mathbf{A}) \neq 0$ allora la matrice è invertibile e precisamente

$$(\mathbf{A}^{-1})_{jk} = \mathfrak{A}_{kj} / \det(\mathbf{A}) .$$

🔗 La formula di Lagrange per l'inversa di una matrice è consigliabile solo per matrici di piccole dimensioni; infatti il numero di operazioni necessarie cresce come il fattoriale della dimensione della matrice. Esistono algoritmi molto più efficienti (Gauss) e questi sono utilizzati nei programmi di calcolo automatico. Si veda [GL96]. ◀

PROBLEMA 3-5. Dimostrare che la matrice inversa dell'operatore

$$\mathcal{A} + \varepsilon |\xi\rangle \langle \eta|$$

ottenuto aggiungendo un proiettore all'operatore invertibile \mathcal{A} è esprimibile, per ε sufficientemente piccolo, come

$$(\mathcal{A} + \varepsilon |\xi\rangle \langle \eta|)^{-1} = \mathcal{A}^{-1} - \varepsilon (1 + \varepsilon \langle \eta | \mathcal{A}^{-1} | \xi \rangle)^{-1} \mathcal{A}^{-1} |\xi\rangle \langle \eta | \mathcal{A}^{-1} .$$

Nel caso più semplice, la formula ci dà

$$(\mathbb{1} + \varepsilon |\xi\rangle \langle \eta|)^{-1} = \mathbb{1} - \varepsilon (1 + \varepsilon \langle \eta | \xi \rangle)^{-1} |\xi\rangle \langle \eta| .$$

che si ricava facilmente sviluppando in serie di Taylor.

3.2. Teoria spettrale

Veniamo ora al punto centrale della teoria degli operatori lineari. Si presenta spesso il problema di trattare funzioni di operatori lineari, in genere funzioni analitiche*. Per superare le limitazioni inerenti alla definizione attraverso serie di potenze, discuteremo ora la possibilità di stabilire una formula di tipo Cauchy

$$(3.4) \quad \hat{f}(\mathcal{A}) \equiv \frac{1}{2\pi i} \oint_{\Gamma} (z\mathbb{1} - \mathcal{A})^{-1} f(z) dz$$

cercando un'opportuno cammino di integrazione Γ in campo complesso. La notazione \hat{f} sta a distinguere tra l'applicazione $f: \mathbb{C} \mapsto \mathbb{C}$ e $\hat{f}: \mathfrak{B} \mapsto \mathfrak{B}$. Se $f(z)$ non è intera ma è definita analitica solo in un dominio $\mathcal{D} \subset \mathbb{C}$, allora sarà necessario delimitare anche un dominio di definizione per \hat{f} . Abbiamo visto nel capitolo precedente i criteri per stabilire l'esistenza dell'operatore inverso

$$\mathcal{R}(z) = (z\mathbb{1} - \mathcal{A})^{-1}$$

che chiameremo d'ora in poi l'*operatore risolvente* di \mathcal{A} . Definiamo ora lo *spettro* di \mathcal{A} , indicandolo con $\sigma(\mathcal{A})$, l'insieme dei valori complessi di z per i quali $z\mathbb{1} - \mathcal{A}$ **non** è invertibile. Notiamo che, per definizione, si ha che per ogni $\lambda \in \sigma(\mathcal{A})$ deve esistere almeno un vettore $|\lambda\rangle$ diverso dal vettore nullo tale che

$$(\lambda\mathbb{1} - \mathcal{A}) |\lambda\rangle = 0$$

ossia

$$\mathcal{A} |\lambda\rangle = \lambda |\lambda\rangle .$$

Il vettore $|\lambda\rangle$ ha dunque la proprietà di essere lasciato *invariato in direzione* dall'applicazione di \mathcal{A} ; si dice che $|\lambda\rangle$ è un *autovettore* di \mathcal{A} *appartenente all'autovalore* $\lambda \in \sigma(\mathcal{A})$.

Teor.3.2.1

TEOREMA 3.2.1. *Lo spettro di un qualunque operatore lineare \mathcal{A} in uno spazio finito dimensionale è un insieme finito non-vuoto di punti nel piano complesso.*

PROVA. La dimostrazione si dà facilmente introducendo una base e la rappresentazione matriciale di $\mathcal{A} \leftarrow \|\mathbf{a}_{ij}\|$. Allora la condizione $\lambda \in \sigma(\mathcal{A})$ si traduce in

$$\Delta(\lambda) \equiv \det(\lambda\delta_{ij} - \mathbf{a}_{ij}) = 0$$

Ora, è immediato verificare che $\Delta(\lambda)$ è un polinomio di grado $\mathfrak{n} = \dim(\mathfrak{E})$ in λ ; ne discende che l'equazione $\Delta(\lambda) = 0$ ammette \mathfrak{n} radici complesse, eventualmente non tutte distinte. ■

✎ È immediato verificare che Δ non dipende dalla scelta della base. Infatti abbiamo visto che le matrici che rappresentano lo stesso operatore in due basi differenti sono legate tra loro da una relazione di equivalenza $\mathbf{A}' = \mathbf{S} \mathbf{A} \mathbf{S}^{-1}$: ne segue

$$\det(\lambda\mathbb{1} - \mathbf{A}') = \det(\lambda\mathbb{1} - \mathbf{S} \mathbf{A} \mathbf{S}^{-1}) = \det(\mathbf{S} (\lambda\mathbb{1} - \mathbf{A}) \mathbf{S}^{-1}) = \det(\lambda\mathbb{1} - \mathbf{A})$$

Dato che Δ dipende solo dall'operatore \mathcal{A} e non dalla base scelta, esso viene chiamato polinomio caratteristico di \mathcal{A} . ◀

*In seguito considereremo anche funzioni più generali quali $|\mathcal{X}|$, $\text{sgn}(\mathcal{X})$, $\arg(\mathcal{X})$, etc.

Siamo ora in grado di formulare una definizione di funzione analitica di operatore in termini di una formula integrale simile a quella di Cauchy (formula di Riesz–Dunford [GL74])

DEF. 3.2.1. Sia \mathcal{A} un qualunque operatore lineare in $\mathfrak{B}(\mathfrak{C})$; sia $\sigma(\mathcal{A})$ il suo spettro e sia $f(z)$ una funzione analitica definita regolare in un dominio \mathcal{D} contenente un aperto (non necessariamente connesso) $\mathcal{U} \subset \mathbb{C}$ di cui gli autovalori siano punti interni; inoltre supponiamo che la frontiera $\partial\mathcal{U}$ di \mathcal{U} sia una curva di Jordan orientata nel modo usuale (antiorario). Si definisce allora

Def.3.2.1

(3.5)

$$\hat{f}(\mathcal{A}) \equiv \frac{1}{2\pi i} \oint_{\partial\mathcal{U}} (z\mathbb{I} - \mathcal{A})^{-1} f(z) dz$$

ESEMPIO 3-2. Sia $\mathcal{A} \rightarrow \sigma_1$, $f(z) = e^{itz}$. Allora

$$\begin{aligned} \mathcal{R}(z) &= (z\mathbb{I} - \sigma_1)^{-1} = \begin{pmatrix} z & -1 \\ -1 & z \end{pmatrix}^{-1} \\ &= (z^2 - 1)^{-1} \begin{pmatrix} z & 1 \\ 1 & z \end{pmatrix} \end{aligned}$$

Se ne deduce che $\sigma(\mathcal{A}) = \{1, -1\}$. Come dominio \mathcal{U} possiamo scegliere ad esempio un cerchio di raggio maggiore di uno centrato in $z = 0$. Si trova allora

$$\widehat{\exp}(\mathcal{A}) \leftarrow \frac{1}{2\pi i} \oint_{|z|=2} \frac{e^{itz}}{z^2 - 1} \begin{pmatrix} z & 1 \\ 1 & z \end{pmatrix} dz$$

e il nostro compito si limita al calcolo dei due integrali

$$\begin{aligned} I_1 &= \frac{1}{2\pi i} \oint_{|z|=2} \frac{e^{itz}}{z^2 - 1} dz \\ I_2 &= \frac{1}{2\pi i} \oint_{|z|=2} \frac{e^{itz}}{z^2 - 1} z dz \end{aligned}$$

che si valutano con il metodo dei residui. Il risultato finale è stesso che avremmo ottenuto sommando la serie di Taylor

$$\widehat{\exp}(it\sigma_1/2) = \begin{pmatrix} \cos \frac{1}{2}t & i \sin \frac{1}{2}t \\ i \sin \frac{1}{2}t & \cos \frac{1}{2}t \end{pmatrix}.$$



PROBLEMA 3-6. Applicare la definizione al caso $f(z) = e^{itz}$ e

$$\mathcal{A} \leftarrow \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$



Una prima proprietà importante della applicazione \hat{f} è la seguente

Teor.3.2.2TEOREMA 3.2.2. *Vale la relazione*

$$\widehat{f}_1(\mathcal{A}) \widehat{f}_2(\mathcal{A}) = \widehat{f_1 f_2}(\mathcal{A})$$

se f_1 e f_2 sono definite in un dominio comune che soddisfi alle ipotesi della Def. 3.2.1.

Dimostriamo preliminarmente il

LEMMA. *Per il risolvete vale l'identità*

$$(3.6) \quad \mathcal{R}(z) \mathcal{R}(w) = - \frac{\mathcal{R}(z) - \mathcal{R}(w)}{z - w}$$

PROVA. La prova del lemma è ovvia: si moltiplicano ambo i membri per $(z\mathbb{1} - \mathcal{A})(w\mathbb{1} - \mathcal{A})$. Veniamo ora al teorema. Scegliamo il cammino di integrazione $\partial\mathcal{U}_1$ interamente all'interno di $\partial\mathcal{U}_2$. Si ha allora

$$\widehat{f}_1(\mathcal{A}) \widehat{f}_2(\mathcal{A}) = \frac{1}{(2\pi i)^2} \oint_{\partial\mathcal{U}_1} \oint_{\partial\mathcal{U}_2} f_1(z) f_2(z') \mathcal{R}(z) \mathcal{R}(z') dz dz'$$

Utilizzando il Lemma si ottiene

$$\begin{aligned} \widehat{f}_1(\mathcal{A}) \widehat{f}_2(\mathcal{A}) &= - \frac{1}{(2\pi i)^2} \oint_{\partial\mathcal{U}_1} f_1(z) \mathcal{R}(z) \underbrace{\oint_{\partial\mathcal{U}_2} \frac{f_2(z')}{z - z'} dz'}_{\Rightarrow -2\pi i f_2(z)} dz \\ &\quad + \frac{1}{(2\pi i)^2} \oint_{\partial\mathcal{U}_2} f_2(z') \mathcal{R}(z') \underbrace{\oint_{\partial\mathcal{U}_1} \frac{f_1(z)}{z - z'} dz}_{\Rightarrow 0} dz' \\ &= \frac{1}{2\pi i} \oint_{\partial\mathcal{U}_1} f_1(z) f_2(z) \mathcal{R}(z) dz \end{aligned}$$

avendo applicato due volte il teorema di Cauchy, nel secondo caso z è esterno al cammino di integrazione e quindi l'integrale si annulla. ■

Ciò mostra in definitiva che la definizione di \widehat{f} rispetta la struttura algebrica delle funzioni analitiche sotto somma e prodotto. Questo fatto sarà molto utile nel seguito.

3.2.1. Proprietà dello spettro.**Teor.3.2.3**

TEOREMA 3.2.3. *Autovettori corrispondenti ad autovalori distinti sono fra loro linearmente indipendenti.*

PROVA. Siano $\{\lambda_1, \dots, \lambda_r\}$ autovalori distinti di \mathcal{A} e $\{|\lambda_1\rangle, \dots, |\lambda_r\rangle\}$ gli autovettori corrispondenti. Supponiamo che valga la relazione

$$\sum_{k=1}^r c_k |\lambda_k\rangle = 0$$

Applicando $1, 2, \dots, r-1$ volte l'operatore \mathcal{A} si otterrà

$$\sum_1^r c_k \lambda_k^{j-1} |\lambda_k\rangle = \mathbf{0} \quad (j = 1, \dots, r) .$$

Ma la matrice $\|\lambda_k^{j-1}\|$, nota come matrice di Vandermonde ha determinante uguale a

$$\prod_{j < k} (\lambda_j - \lambda_k)$$

ed è diverso da zero. La matrice è dunque invertibile e si conclude che tutte le costanti c_k devono annullarsi. ■

COROLLARIO. *Se un operatore lineare ammette tutti autovalori distinti in numero pari alla dimensione dello spazio i suoi autovettori formano una base.*

Un operatore lineare i cui autovettori formino una base è detto *diagonalizzabile*, nel senso che la sua matrice rappresentativa nella base degli autovettori assume la forma

$$\mathcal{A} \leftrightarrow \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{n-1} & 0 \\ 0 & \dots & 0 & 0 & \lambda_n \end{pmatrix}$$

ossia, in breve, $a_{ij} = \lambda_j \delta_{ij}$. La formula per $\hat{f}(\mathcal{A})$, nel caso di matrici diagonali, dà semplicemente

$$\hat{f}(\mathcal{A}) \leftrightarrow \begin{pmatrix} f(\lambda_1) & 0 & \dots & \dots & 0 \\ 0 & f(\lambda_2) & 0 & \dots & 0 \\ 0 & 0 & f(\lambda_3) & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & f(\lambda_{n-1}) & 0 \\ 0 & \dots & 0 & 0 & f(\lambda_n) \end{pmatrix}$$

e dunque l'operazione di “diagonalizzare” una matrice risolve completamente il problema di calcolo per $\hat{f}(\mathcal{A})$. Però *non tutti gli operatori lineari* sono diagonalizzabili. Esamineremo nel capitolo seguente il caso più generale e troveremo le condizioni necessarie e sufficienti per la diagonalizzabilità.

3.2.2. La decomposizione spettrale. Siano $\{\lambda_1, \lambda_2, \dots, \lambda_m\}$ gli autovalori distinti di \mathcal{A} ciascuno con molteplicità $\{r_1, r_2, \dots, r_m\}$ ($\sum r_k = n = \dim(\mathfrak{E})$). Ciò equivale ad affermare che il polinomio caratteristico è dato da

$$\Delta(\mathcal{A}, z) = \det(z\mathbb{1} - \mathcal{A}) = \prod_{k=1}^m (z - \lambda_k)^{r_k} .$$

Il generico elemento di matrice del risolvete è della forma

$$(z\mathbb{1} - \mathcal{A})_{ij} = \frac{\mathfrak{A}_{ij}}{\Delta(\mathcal{A}, z)}$$

ossia è una funzione razionale di z avente poli di ordine $\leq r_k$ nei punti che costituiscono lo spettro. La formula di Riesz–Dunford si può allora trasformare introducendo lo sviluppo di

Taylor di $f(z)$ attorno a ciascun autovalore. Siano Γ_k contorni semplici ciascuno contenente al suo interno un solo autovalore (vedi Fig. 3-3).

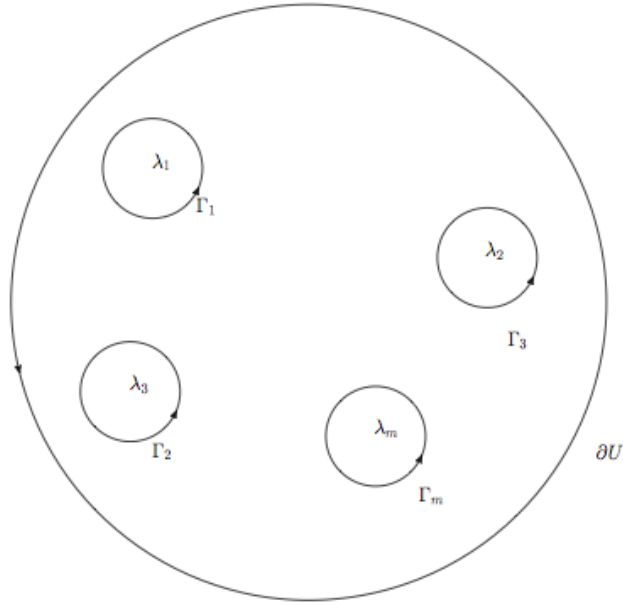


FIGURA 3-3. $\oint_{\partial U} = \sum_k \oint_{\Gamma_k}$

Troviamo dunque

$$\begin{aligned}
 (3.7) \quad \hat{f}(\mathcal{A}) &= \frac{1}{2\pi i} \sum_{k=1}^m \oint_{\Gamma_k} (z\mathbb{1} - \mathcal{A})^{-1} \sum_{\ell=0}^{r_k-1} \frac{f^{(\ell)}(\lambda_k)}{\ell!} (z - \lambda_k)^\ell dz \\
 &\equiv \sum_{k=1}^m \sum_{\ell=0}^{r_k-1} \frac{f^{(\ell)}(\lambda_k)}{\ell!} \mathbb{E}_k^{(\ell)},
 \end{aligned}$$

dove si sono introdotti gli operatori $\mathbb{E}_k^{(\ell)}$ secondo la definizione

$$\mathbb{E}_k^{(\ell)} \equiv \frac{1}{2\pi i} \oint_{\Gamma_k} (z\mathbb{1} - \mathcal{A})^{-1} (z - \lambda_k)^\ell dz.$$

Scriveremo allora

$$(3.8) \quad \hat{f}(\mathcal{A}) = \sum_{\lambda \in \sigma(\mathcal{A})} \sum_{\ell=0}^{r(\lambda)-1} \frac{f^{(\ell)}(\lambda)}{\ell!} \mathbb{E}_\lambda^{(\ell)},$$

Siamo ora in grado di dimostrare alcune proprietà fondamentali dalla corrispondenza $f(z) \leftrightarrow \hat{f}(\mathcal{A})$ definita attraverso la formula di Riesz–Dunford - una proprietà la abbiamo già mostrata in precedenza (Teor.3.2.2).

Teor.3.2.4

TEOREMA 3.2.4. Se $f(z) = z^m$ allora $\hat{f}(\mathcal{A}) = \mathcal{A}^m$

PROVA. Segue dal Teor.3.2.2, tuttavia una dimostrazione diretta offre qualche spunto interessante (Cauchy al lavoro). Per $\mathbf{m} = 0$ si ha

$$\frac{1}{2\pi i} \oint_{\partial \mathcal{U}} (z\mathbb{1} - \mathcal{A})^{-1} dz = \frac{1}{2\pi i} \oint_{\partial \mathcal{U}} (\mathbb{1} - \mathcal{A}/z)^{-1} \frac{dz}{z}$$

Scegliendo il contorno come un cerchio centrato in zero e raggio R , l'integrale

$$\oint_{|z|=R} (\mathbb{1} - \mathcal{A}/R e^{-i\theta})^{-1} d\theta$$

non dipende da R e quindi possiamo prendere R grande a piacere. Nel limite si ha

$$\mathbb{1} = \frac{1}{2\pi i} \oint (z\mathbb{1} - \mathcal{A})^{-1} dz \equiv \sum_{\lambda \in \sigma(\mathcal{A})} \mathbb{E}_{\lambda}^{(0)}$$

Analogamente si trova

$$\mathcal{A} = \sum_{\lambda \in \sigma(\mathcal{A})} \left(\lambda \mathbb{E}_{\lambda}^{(0)} + \mathbb{E}_{\lambda}^{(1)} \right)$$

e in generale

$$\mathcal{A}^k = \sum_{\lambda \in \sigma(\mathcal{A})} \left(\lambda^k \mathbb{E}_{\lambda}^{(0)} + k \lambda^{k-1} \mathbb{E}_{\lambda}^{(1)} + \dots \right)$$

che ci porta all'identità, valida per ogni polinomio $p(z)$,

$$p(\mathcal{A}) = \sum_{\lambda \in \sigma(\mathcal{A})} \sum_{\ell=0}^{r_{\lambda}-1} \frac{p^{(\ell)}(\lambda)}{\ell!} \mathbb{E}_{\lambda}^{(\ell)}.$$

■

⚡ Se mandiamo all'infinito il grado del polinomio e la serie così ottenuta converge in un cerchio che contiene al suo interno lo spettro di \mathcal{A} , allora la formula precedente mostra che anche la serie $\sum_k c_k \mathcal{A}^k$ converge. La definizione che abbiamo dato in termini della formula di Riesz–Dunford è però più generale in quanto dà significato a $\hat{f}(\mathcal{A})$ anche se non esiste alcuno sviluppo in serie convergente. Un esempio elementare è il seguente: si voglia calcolare $\widehat{\log}(\mathcal{A})$ e lo spettro contenga un punto λ e il suo opposto $-\lambda$. Ogni cerchio che contenga entrambi deve contenere anche il punto di diramazione del logaritmo e perciò non esiste sviluppo in serie che si presti alla definizione, mentre l'integrale di Riesz–Dunford è perfettamente definito – pur richiedendo una scelta convenzionale della determinazione del logaritmo. ◀

Dimostriamo ora una proprietà notevole degli operatori \mathbb{E}_{λ} .

TEOREMA 3.2.5. $\mathbb{E}_{\lambda}^{(\ell)} \mathbb{E}_{\lambda'}^{(\ell')} = \delta_{\lambda\lambda'} \mathbb{E}_{\lambda}^{(\ell+\ell')}$

Teor.3.2.5

PROVA. Utilizziamo ancora il Lemma 3.2. Si trova

$$\begin{aligned}
 (3.9) \quad \mathbb{E}_\lambda^{(\ell)} \mathbb{E}_{\lambda'}^{(\ell')} &= \frac{1}{(2\pi i)^2} \oint_{\Gamma} \oint_{\Gamma'} (z - \lambda)^\ell (z - \lambda')^{\ell'} \mathcal{R}(z) \mathcal{R}(z') dz dz' \\
 &= \frac{1}{(2\pi i)^2} \oint_{\Gamma} dz \mathcal{R}(z) (z - \lambda)^\ell \underbrace{\oint_{\Gamma'} (z - \lambda')^{\ell'} (z' - z)^{-1} dz'}_{(*)} \\
 &\quad + \frac{1}{(2\pi i)^2} \oint_{\Gamma'} dz' \mathcal{R}(z') (z' - \lambda')^{\ell'} \underbrace{\oint_{\Gamma} (z - \lambda)^\ell (z - z')^{-1} dz}_{(**)}
 \end{aligned}$$

Consideriamo prima il caso $\lambda \neq \lambda'$: in entrambi gli integrali che non contengono il risolvente ($*$ e $**$) la singolarità è esterna al cammino di integrazione e dunque otteniamo il fattore $\delta_{\lambda\lambda'}$. Per $\lambda = \lambda'$, scegliamo Γ interno a Γ' . Allora $(**)$ si annulla e il primo integrale ci dà quanto affermato dal teorema. ■

PROBLEMA 3-7. Ricavare il Teor.3.2.2 utilizzando il Teor.3.2.5

Per ogni $\lambda \in \sigma(\mathcal{A})$ sappiamo che $\mathbb{E}_\lambda^{(\ell)} = 0$ per $\ell \geq r_\lambda$ (ricordiamo che r_λ è la molteplicità dell'autovalore). D'altronde, per il teorema appena dimostrato, abbiamo che $\mathbb{E}_\lambda^{(\ell)} = (\mathbb{E}_\lambda^{(1)})^\ell$, ossia tutti gli operatori \mathbb{E} sono esprimibili come potenze del caso $\ell = 1$. Conviene allora introdurre una notazione che alleggerisce le formule:

$$\text{DEF. 3.2.2. } \mathbb{E}_\lambda^{(0)} \equiv \mathbb{P}_\lambda, \quad \mathbb{E}_\lambda^{(1)} \equiv \mathbb{E}_\lambda.$$

Potremo dunque riscrivere lo sviluppo (3.8) nella forma

$$(3.10) \quad \hat{f}(\mathcal{A}) = \sum_{\lambda \in \sigma(\mathcal{A})} \left(f(\lambda) \mathbb{P}_\lambda + \sum_{\ell=1}^{r_\lambda-1} \frac{f^{(\ell)}(\lambda)}{\ell!} \mathbb{E}_\lambda^\ell \right).$$

Come casi particolari notiamo la *relazione di completezza*

$$(3.11) \quad \boxed{\mathbb{1} = \sum_{\lambda \in \sigma(\mathcal{A})} \mathbb{P}_\lambda}$$

e la *decomposizione spettrale*

$$(3.12) \quad \boxed{\mathcal{A} = \sum_{\lambda \in \sigma(\mathcal{A})} (\lambda \mathbb{P}_\lambda + \mathbb{E}_\lambda)}$$

Gli operatori \mathbb{P} e \mathbb{E} soddisfano le identità (già dimostrate)

$$\boxed{\mathbb{P}_\lambda \mathbb{P}_\mu = \delta_{\lambda\mu} \mathbb{P}_\lambda \quad \mathbb{P}_\lambda \mathbb{E}_\mu = \mathbb{E}_\mu \mathbb{P}_\lambda = \delta_{\lambda\mu} \mathbb{E}_\lambda}$$

⚡ Operatori, che come gli operatori \mathbb{P} , soddisfano l'identità $\mathbb{P}^2 = \mathbb{P}$ si dicono operatori di proiezione. ◀

DEF. 3.2.3. Sia ν_λ il più piccolo intero ($0 \leq \nu_\lambda \leq r_\lambda - 1$) tale che $\mathbb{E}_\lambda^{\nu_\lambda+1} = 0$; ν_λ è detto l'indice dell'autovalore λ .

Def.3.2.3

TEOREMA 3.2.6. Si dice polinomio minimale di \mathcal{A} il polinomio definito da

Teor.3.2.6

$$\Delta_0(z) = \prod_{\lambda \in \sigma(\mathcal{A})} (z - \lambda)^{1+\nu_\lambda}.$$

Allora vale $\Delta_0(\mathcal{A}) = 0$, e Δ_0 è il polinomio di grado minimo tra tutti quelli per cui $\mathbf{p}(\mathcal{A}) = 0$.

PROVA. Si applichi la (3.10) con $\mathbf{f}(z) = \Delta_0(z)$. Ovviamente un polinomio di grado inferiore darebbe luogo a una derivata non nulla in qualche termine del tipo $\frac{f^{(\ell)}(\lambda)}{\ell!} \mathbb{E}_\lambda^\ell$. Si noti che il polinomio caratteristico soddisfa $\Delta(\mathcal{A}) = 0$, un risultato che era noto a Cayley.

■

Siamo ora in grado di formulare il teorema generale sulla diagonalizzabilità.

TEOREMA 3.2.7. Per ogni operatore lineare \mathcal{A} in uno spazio lineare n -dimensionale \mathfrak{E} , le seguenti affermazioni sono equivalenti:

Teor.3.2.7

- i) \mathcal{A} è diagonalizzabile, cioè ammette una base di autovettori;
- ii) $\mathbb{E}_\lambda = 0$ per ogni $\lambda \in \sigma(\mathcal{A})$;
- iii) tutti gli elementi di matrice del risolvente $\mathcal{R}(z) = (z\mathbb{1} - \mathcal{A})^{-1}$ presentano al più poli del primo ordine;
- iv) \mathcal{A} soddisfa un'equazione algebrica con tutte radici semplici.

PROVA. Dimostriamo (i) \implies (iv) \implies (iii) \implies (ii) \implies (i). Nella base degli autovettori \mathcal{A} è rappresentato da una matrice diagonale

$$\mathcal{A} \leftarrow \begin{pmatrix} \lambda_1 \mathbb{1}_{r_{\lambda_1}} & 0 & \dots & 0 \\ 0 & \lambda_2 \mathbb{1}_{r_{\lambda_2}} & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 \mathbb{1}_{r_{\lambda_3}} & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{m-1} \mathbb{1}_{r_{\lambda_{m-1}}} & 0 \\ 0 & \dots & 0 & 0 & \lambda_m \mathbb{1}_{r_{\lambda_m}} \end{pmatrix}$$

dove $\mathbb{1}_k$ è la matrice unità $k \times k$; ne segue che il prodotto $(\mathcal{A} - \lambda_1 \mathbb{1})(\mathcal{A} - \lambda_2 \mathbb{1}) \dots (\mathcal{A} - \lambda_m \mathbb{1})$ risulta nullo in quanto ogni singolo fattore presenta un blocco nullo ($\mathbb{O}_{r_{\lambda_k}}$) lungo la diagonale ossia

$$\mathcal{A} - \lambda_1 \mathbb{1} \leftarrow \begin{pmatrix} \mathbb{O}_{r_{\lambda_1}} & 0 & \dots & 0 \\ 0 & (\lambda_2 - \lambda_1) \mathbb{1}_{r_{\lambda_2}} & 0 & \dots & 0 \\ 0 & 0 & (\lambda_3 - \lambda_1) \mathbb{1}_{r_{\lambda_3}} & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & (\lambda_{m-1} - \lambda_1) \mathbb{1}_{r_{\lambda_{m-1}}} & 0 \\ 0 & \dots & 0 & 0 & (\lambda_m - \lambda_1) \mathbb{1}_{r_{\lambda_m}} \end{pmatrix}$$

... per tutti i blocchi diagonali fino all' m -esimo ...

$$\mathcal{A} - \lambda_m \mathbb{1} \leftrightarrow \begin{pmatrix} (\lambda_1 - \lambda_m) \mathbb{1}_{r_{\lambda_1}} & 0 & \dots\dots\dots & 0 \\ 0 & (\lambda_2 - \lambda_m) \mathbb{1}_{r_{\lambda_2}} & 0 & \dots\dots\dots & 0 \\ 0 & 0 & (\lambda_3 - \lambda_m) \mathbb{1}_{r_{\lambda_3}} & 0 & \dots\dots\dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots\dots\dots & 0 & (\lambda_{m-1} - \lambda_m) \mathbb{1}_{r_{\lambda_{m-1}}} & 0 \\ 0 & \dots\dots\dots & & 0 & \mathbb{O}_{r_{\lambda_m}} \end{pmatrix}$$

e questo fissa (i) \implies (iv). Sia ora $\Phi(\mathcal{A}) \equiv \prod_k (\mathcal{A} - z_k) = 0$, con tutti gli z_k distinti (iv)). Ne segue

$$\mathcal{R}(z) = (z\mathbb{1} - \mathcal{A})^{-1} \equiv \frac{1}{\Phi(z)} \frac{\Phi(z)\mathbb{1} - \Phi(\mathcal{A})}{z\mathbb{1} - \mathcal{A}};$$

Ora $(\Phi(z)\mathbb{1} - \Phi(\mathcal{A})) / (z\mathbb{1} - \mathcal{A})$ è un polinomio in z e quindi le uniche singolarità del risolvante vengono da $\Phi(z)$ e dunque si tratta di poli semplici ((iv) \implies (iii)). (iii) \implies (ii) segue immediatamente dalla definizione di \mathbb{E}_λ . Infine da (ii) segue che la decomposizione spettrale assume la forma più semplice

$$\boxed{\mathcal{A} = \sum_{\lambda \in \sigma(\mathcal{A})} \lambda \mathbb{P}_\lambda}$$

Dalla relazione di completezza Eq. (3.11) inoltre per ogni vettore vale la decomposizione

$$|\xi\rangle = \sum_{\lambda \in \sigma(\mathcal{A})} \mathbb{P}_\lambda |\xi\rangle.$$

Ma ogni componente $\mathbb{P}_\lambda |\xi\rangle$ soddisfa l'equazione agli autovalori

$$\mathcal{A} \mathbb{P}_\lambda |\xi\rangle = \sum_{\mu} \mu \mathbb{P}_\mu \mathbb{P}_\lambda |\xi\rangle = \lambda \mathbb{P}_\lambda |\xi\rangle.$$

il che mostra che ogni vettore è decomponibile in componenti che sono autovettori di \mathcal{A} . Assumiamo ora che gli autovettori di \mathcal{A} *non formino una base*. Sia **|altro>** uno dei vettori linearmente indipendente dagli autovettori che si debbono aggiungere per ottenere una base. Ma anche **|altro>** si può decomporre in $|\text{altro}\rangle = \sum_{\lambda \in \sigma(\mathcal{A})} \mathbb{P}_\lambda |\text{altro}\rangle$ ed è quindi esprimibile come combinazione lineare di autovettori, il che contraddice l'ipotesi. ■

◀ Gli operatori di proiezione ($X^2 = X$) sono diagonalizzabili; al contrario gli operatori nilpotenti ($\mathbb{E}^{v+1} = 0$) non sono diagonalizzabili. ▶

A titolo di curiosità, perchè la cosa non rappresenta grandi vantaggi sul piano pratico, citiamo il seguente risultato: per ogni funzione $f(z)$ tale che $\hat{f}(\mathcal{A})$ sia ben definita, esiste un polinomio $P_f(\mathcal{A})$ di grado non superiore a $\dim(\mathcal{C}) - 1$ tale che $P_f(\mathcal{A}) \equiv \hat{f}(\mathcal{A})$. Il polinomio è dato dalla formula integrale

$$P_f(\zeta) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{\Delta_0(z)} \frac{\Delta_0(z)\mathbb{1} - \Delta_0(\zeta)}{z - \zeta} dz.$$

infatti

$$P_f(\mathcal{A}) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{\Delta_0(z)} \frac{\Delta_0(z)\mathbb{1} - \Delta_0(\mathcal{A})}{z\mathbb{1} - \mathcal{A}} dz$$

e la formula coincide con la definizione di $\hat{f}(\mathcal{A})$ in quanto $\Delta_0(\mathcal{A}) \equiv \mathbb{O}$.

✎ Ci si può chiedere cosa succede nel caso in cui non valgano le condizioni del Teor.3.2.7. L'operatore in questo caso non possiede abbastanza autovettori per costruire una base. La sua forma matriciale standard non sarà una matrice diagonale, bensì una matrice del tipo

$$\begin{pmatrix} \lambda_1 \mathbb{1}_{r_{\lambda_1}} + \mathbb{E}_{r_{\lambda_1}} & 0 & \dots & 0 \\ 0 & \lambda_2 \mathbb{1}_{r_{\lambda_2}} + \mathbb{E}_{r_{\lambda_2}} & \dots & 0 \\ 0 & 0 & \lambda_3 \mathbb{1}_{r_{\lambda_3}} + \mathbb{E}_{r_{\lambda_3}} & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{m-1} \mathbb{1}_{r_{\lambda_{m-1}}} + \mathbb{E}_{r_{\lambda_{m-1}}} \\ 0 & \dots & 0 & \lambda_m \mathbb{1}_{r_{\lambda_m}} + \mathbb{E}_{r_{\lambda_m}} \end{pmatrix}$$

dove le matrici \mathbb{E} rappresentano gli operatori \mathbb{E}_{λ} ; si tratta di matrici nilpotenti, caratterizzate cioè dal fatto che la loro potenza $\mathbb{E}_{\lambda}^{v_{\lambda}+1}$ si annulla. In generale si ha che il difetto nel numero di autovettori per arrivare a n è dato dalla somma degli indici v_{λ} . L'analisi del caso non-diagonalizzabile è chiaramente più complessa e a consolazione dello studioso, le applicazioni alla Meccanica Quantistica sono per lo più libere da questa complicazione.

Un esempio di matrice di questo tipo è la seguente

$$X = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha \end{pmatrix}$$

Il polinomio caratteristico è $\Delta(z) = z^3(z - \alpha)$, e la matrice ammette solo due autovettori

$$|1\rangle \leftarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |2\rangle \leftarrow \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

L'autovalore 0 ha molteplicità 3, ma il risolvente ha in 0 un polo del terzo ordine:

$$\mathcal{R}(z) = \begin{pmatrix} z^{-1} & z^{-2} & z^{-3} & 0 \\ 0 & z^{-1} & z^{-2} & 0 \\ 0 & 0 & z^{-1} & 0 \\ 0 & 0 & 0 & (z-\alpha)^{-1} \end{pmatrix}$$

e dunque $v_0 = 2, v_{\alpha} = 0$. Si noti che matrici di questo tipo costituiscono un problema anche dal punto di vista numerico, in quanto il calcolo dello spettro può presentare instabilità. Consideriamo infatti una matrice leggermente modificata

$$X(\varepsilon) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \varepsilon & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha \end{pmatrix}$$

Lo spettro in questo caso è dato da $\sigma(X(\varepsilon)) = \{\varepsilon^{1/3}, \varepsilon^{1/3} e^{2\pi i/3}, \varepsilon^{1/3} e^{4\pi i/3}, \alpha\}$ cioè esso è costituito da quattro autovalori distinti e la matrice risulta diagonalizzabile. Immaginiamo cosa succede se invece di un blocco 3-dimensionale si trattasse di un blocco K -dimensionale con K molto grande e ammettiamo che ε sia in realtà un piccolo errore che si è inserito nel calcolo della matrice

ottenuta per via numerica. L'errore sullo spettro risulta dell'ordine di $\varepsilon^{1/K}$ che può risultare molto grande anche se $\varepsilon \approx 10^{-16}$. Le problematiche relative alla diagonalizzazione di matrici anche di grandi dimensioni sono affrontate esaurientemente su [GL96]. ◀

3.3. Operatori autoaggiunti, unitari, normali

Tutte le considerazioni riguardanti la decomposizione spettrale di operatori lineari presentate finora hanno fatto riferimento a un generico spazio lineare finito-dimensionale. Studiamo ora il caso di uno spazio unitario e restringiamo l'attenzione a classi particolari di operatori lineari per le quali esiste un teorema di diagonalizzabilità molto semplice.

Definiamo innanzitutto il concetto di *aggiunto Hermitiano* di un operatore lineare (temporaneamente utilizziamo la “vecchia” notazione per il prodotto interno)

Def.3.3.1

DEF. 3.3.1. Dato un operatore lineare H , per ogni vettore $\eta \in \mathfrak{C}$ la relazione

$$\phi_\eta(\xi) = \langle \eta, H\xi \rangle$$

definisce un funzionale lineare $\xi \mapsto \phi_\eta(\xi)$. Per il Teor.2.4.2 esiste un unico vettore ϕ_η^\sharp tale che $\phi_\eta(\xi) = \langle \phi_\eta^\sharp, \xi \rangle$. Si pone allora

$$\phi_\eta^\sharp = H^\dagger \eta.$$

Si verifica facilmente che H^\dagger è un operatore lineare; esso è detto l'*aggiunto Hermitiano* di H . Vale dunque la relazione $\langle \eta, H\xi \rangle = \langle H^\dagger \eta, \xi \rangle$ ovvero nel formalismo di Dirac

$$\langle \eta | H | \xi \rangle = \overline{\langle \xi | H^\dagger | \eta \rangle}.$$

In una base ortonormale, la matrice che rappresenta H^\dagger è data dalla coniugata Hermitiana di H :

$$\langle i | H^\dagger | j \rangle = \overline{\langle j | H | i \rangle}.$$

Diamo ora alcune definizioni di fondamentale importanza in vista delle applicazioni:

Def.3.3.2

DEF. 3.3.2. Un operatore lineare \mathcal{H} si dice *autoaggiunto* se $\boxed{\mathcal{H} = \mathcal{H}^\dagger}$.

Def.3.3.3

DEF. 3.3.3. Un operatore lineare \mathcal{U} si dice *unitario* se $\boxed{\mathcal{U}\mathcal{U}^\dagger = \mathcal{U}^\dagger\mathcal{U} = \mathbb{1}}$.

Def.3.3.4

DEF. 3.3.4. Un operatore lineare \mathcal{N} si dice *normale* se $\boxed{\mathcal{N}\mathcal{N}^\dagger = \mathcal{N}^\dagger\mathcal{N}}$.

Illustreremo ora alcune proprietà notevoli degli operatori così definiti.

Teor.3.3.1

TEOREMA 3.3.1. Tutti gli autovalori di un operatore autoaggiunto sono reali.

PROVA.

$$H|\lambda\rangle = \lambda|\lambda\rangle \implies \langle \lambda | H | \lambda \rangle = \lambda \langle \lambda | \lambda \rangle \implies \overline{\lambda} \langle \lambda | \lambda \rangle = \overline{\langle \lambda | H | \lambda \rangle} = \langle \lambda | H^\dagger | \lambda \rangle = \lambda \langle \lambda | \lambda \rangle$$

■

LEMMA. Sia $|\lambda\rangle$ un autovettore dell'operatore autoaggiunto H , appartenente all'autovalore λ . L'insieme dei vettori ortogonali a $|\lambda\rangle$ è un sottospazio di dimensione $n-1$ che è lasciato invariante da H (ciò significa che se $\langle \lambda | \xi \rangle = 0$ allora vale anche $\langle \lambda | H | \xi \rangle = 0$).

(La dimostrazione è immediata).

Teor.3.3.2

TEOREMA 3.3.2. Ogni operatore autoaggiunto ammette una base di autovettori tra loro ortogonali.

PROVA. Sappiamo che ogni operatore lineare \mathcal{H} ammette almeno un autovettore. Sia $|\lambda_1\rangle$ tale autovettore; per il lemma, il sottospazio λ_1^\perp ortogonale a $|\lambda_1\rangle$ è invariante sotto l'azione di \mathcal{H} ; dunque \mathcal{H} definisce un operatore autoaggiunto nel sottospazio λ_1^\perp ; esisterà dunque un secondo autovettore $|\lambda_2\rangle \in \lambda_1^\perp$; il sottospazio λ_{12}^\perp ortogonale a entrambi gli autovettori è a sua volta lasciato invariante da \mathcal{H} , e l'argomento si ripete sino ad esaurire tutto lo spazio. ■

n \neq ∞ Nel caso di spazi a infinite dimensioni il quadro è più complesso. Gli autovettori di un operatore autoaggiunto non costituiscono infatti, in generale, una base ortonormale. Si veda il Cap. 4.3. ◀

TEOREMA 3.3.3. *Condizione necessaria e sufficiente affinché un operatore lineare \mathcal{H} sia autoaggiunto è che esista una base ortogonale rispetto a cui \mathcal{H} sia rappresentato da una matrice diagonale con elementi reali.*

Teor.3.3.3

(La dimostrazione è lasciata per esercizio). Ci si può convincere facilmente che ogni operatore autoaggiunto è diagonalizzabile anche attraverso il seguente argomento. Supponiamo che

$$(\mathcal{H} - \lambda \mathbb{1})^m |\lambda\rangle = 0, \quad (m > 1)$$

Segue che

$$(\mathcal{H} - \lambda \mathbb{1})^{m-1} |\lambda\rangle = 0$$

infatti, essendo λ reale si ha

$$\|(\mathcal{H} - \lambda \mathbb{1})^{m-1} |\lambda\rangle\|^2 = \langle \lambda | (\mathcal{H} - \lambda \mathbb{1})^{m-2} (\mathcal{H} - \lambda \mathbb{1})^m |\lambda\rangle = 0$$

Applicando ricorsivamente il risultato si ottiene

$$(\mathcal{H} - \lambda \mathbb{1}) |\lambda\rangle = 0, \quad (m > 1)$$

e dunque il polinomio minimale di ogni operatore autoaggiunto ha tutti fattori semplici e ci troviamo perciò nel caso (iv) del Teor.3.2.7. L'ortogonalità degli autovettori appartenenti ad autovalori distinti è conseguenza immediata della relazione

$$(\lambda - \lambda') \langle \lambda | \lambda' \rangle = \overline{\langle \lambda' | \mathcal{A} | \lambda \rangle} - \langle \lambda | \mathcal{A} | \lambda' \rangle = 0$$

◀ Gli operatori \mathbb{P}_λ nel caso autoaggiunto sono rappresentabili nella forma

$$\mathbb{P}_\lambda = \sum_{\alpha=1}^{r_\lambda} |\lambda, \alpha\rangle \langle \lambda, \alpha|$$

◀

3.3.1. Operatori unitari. Vediamo subito una caratterizzazione geometrica degli operatori definiti nel Cap. 3.3.

TEOREMA 3.3.4. *Un operatore unitario costituisce una trasformazione biunivoca dello spazio in sé che lascia invariato il prodotto interno di ogni coppia di vettori*

Teor.3.3.4

$$\langle \mathcal{U}\eta, \mathcal{U}\xi \rangle = \langle \eta, \xi \rangle.$$

PROVA. La biunivocità segue dal fatto che \mathcal{U} è invertibile, inoltre $\langle \mathcal{U}\eta, \mathcal{U}\xi \rangle = \langle \eta, \mathcal{U}^\dagger \mathcal{U}\xi \rangle = \langle \eta, \xi \rangle.$ ■

Teor.3.3.5

TEOREMA 3.3.5. *Un operatore lineare che lasci invariante il prodotto interno è unitario.*

PROVA. Da $\langle U\eta, U\xi \rangle = \langle \eta, \xi \rangle$ segue $\langle \eta, U^\dagger U \xi \rangle = \langle \eta, \xi \rangle$ e quindi $U^\dagger U = \mathbb{1}$; allora $\det(U^\dagger U) = 1 = |\det(U)|^2$; dunque U è invertibile e l'inverso è U^\dagger . ■

$n \neq \infty$ È da sottolineare il fatto che da $U^\dagger U = \mathbb{1}$ segue $U U^\dagger = \mathbb{1}$ solo nel caso di spazi finito-dimensionali. Nel caso infinito-dimensionale infatti esistono operatori non invertibili che lasciano invariante il prodotto interno. ◀

La matrice che rappresenta un operatore unitario in una qualunque base ortonormale ha proprietà notevoli. Da

$$\sum_k U_{ik} \overline{U_{jk}} = \delta_{ij}$$

segue che le righe (e le colonne) della matrice costituiscono un insieme di n vettori ortogonali e normalizzati.

Teor.3.3.6

TEOREMA 3.3.6. *Gli autovalori di un operatore unitario sono numeri complessi di modulo uno.*

PROVA. Infatti, sia $U|\lambda\rangle = \lambda|\lambda\rangle$, allora

$$\langle \lambda | \lambda \rangle = \langle \lambda | U^\dagger U | \lambda \rangle = |\lambda|^2 \langle \lambda | \lambda \rangle$$

■

Teor.3.3.7

TEOREMA 3.3.7. *Sia $|\lambda\rangle$ un autovettore dell'operatore unitario U . Il sottospazio ortogonale a $|\lambda\rangle$ è lasciato invariante da U . (Ovvio dal momento che U lascia invariato il prodotto interno).*

Lo stesso tipo di argomento applicato al caso degli operatori autoaggiunti porta dunque a concludere che

Teor.3.3.8

TEOREMA 3.3.8. *Ogni operatore unitario è diagonalizzabile in una base ortogonale.*

Abbiamo raggiunto dunque una proprietà simile per tutta la classe degli operatori autoaggiunti e per quella degli operatori unitari. La differenza tra i due casi è costituita dal fatto che lo spettro della prima classe è contenuta sull'asse reale, quello della seconda nella circonferenza di raggio uno. Ciò suggerisce che ogni funzione $f(z)$ che trasformi la retta nel cerchio dovrebbe trasformare un operatore autoaggiunto in uno unitario. Ad es. $f(z) = \exp\{i\alpha z\}$, oppure $f(z) = (z - i)/(z + i)$.

PROBLEMA 3-8. *Dimostrare, utilizzando la decomposizione spettrale, che per ogni operatore autoaggiunto \mathcal{H} , gli operatori $\exp\{i\mathcal{H}\}$ e $(\mathcal{H} - i)(\mathcal{H} + i)^{-1}$ definiscono operatori unitari.*

Operatori autoaggiunti e operatori unitari sono casi particolari di *operatori normali*. Vogliamo ora mostrare che la condizione

$$[\mathcal{N}, \mathcal{N}^\dagger] = \mathcal{N}\mathcal{N}^\dagger - \mathcal{N}^\dagger\mathcal{N} = 0$$

rappresenta la condizione necessaria e sufficiente affinché \mathcal{N} sia diagonalizzabile in una base ortogonale.

3.3.2. Operatori normali.

LEMMA. Siano \mathcal{A} e \mathcal{B} sue operatori commutanti ($\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} \equiv [\mathcal{A}, \mathcal{B}] = 0$). Allora il sottospazio \mathcal{S}_a generato dagli autovettori di \mathcal{A} appartenenti a un certo autovalore a è invariante rispetto a \mathcal{B} . Di conseguenza due operatori che commutano hanno almeno un autovettore in comune.

PROVA. Sia $\mathcal{A}|a\rangle = a|a\rangle$. Allora

$$\mathcal{A}\{\mathcal{B}|a\rangle\} = \mathcal{B}\mathcal{A}|a\rangle = a\{\mathcal{B}|a\rangle\}$$

e dunque $|a\rangle \in \mathcal{S}_a \Rightarrow \mathcal{B}|a\rangle \in \mathcal{S}_a$. \mathcal{B} definisce un operatore lineare nel sottospazio \mathcal{S}_a che contiene tutti gli autovettori di \mathcal{A} appartenenti ad a ed ha perciò almeno un autovettore in \mathcal{S}_a . ■

TEOREMA 3.3.9. Siano \mathcal{A} e \mathcal{B} due operatori autoaggiunti in \mathfrak{E} . Condizione necessaria e sufficiente affinché \mathcal{A} e \mathcal{B} ammettano una base di autovettori comuni è che essi commutino.

Teor.3.3.9

PROVA. Sia $[\mathcal{A}, \mathcal{B}] = 0$. Esiste allora almeno un autovettore comune $|1\rangle$ con $\mathcal{A}|1\rangle = a_1|1\rangle$, $\mathcal{B}|1\rangle = b_1|1\rangle$. Il sottospazio \mathcal{S}_1^\perp ortogonale a $|1\rangle$ è invariante rispetto a entrambi gli operatori, dunque \mathcal{A} e \mathcal{B} definiscono due operatori autoaggiunti commutanti in \mathcal{S}_1^\perp . Sia $|2\rangle$ un autovettore comune in \mathcal{S}_1^\perp . Procedendo successivamente in questo modo, si ottiene una base comune ortogonale. Il viceversa è ovvio, in quanto matrici diagonali commutano tra loro. ■

TEOREMA 3.3.10. Un operatore lineare è diagonalizzabile in una base ortogonale se e solo se esso è normale.

Teor.3.3.10

PROVA. Dimostriamo la necessità. Sia $N_{ij} = \lambda_i \delta_{ij}$ la matrice che rappresenta \mathcal{N} nella base dei suoi autovettori. Allora \mathcal{N}^\dagger è diagonale con elementi $\mathcal{N}_{ij}^\dagger = \bar{\lambda}_i \delta_{ij}$ (per questo è essenziale che la base sia ortogonale) le due matrici rappresentative commutano e dunque \mathcal{N} è normale. Viceversa, supponiamo che \mathcal{N} sia normale, e poniamo

$$\mathcal{H}_1 = \mathcal{N} + \mathcal{N}^\dagger, \quad \mathcal{H}_2 = i(\mathcal{N} - \mathcal{N}^\dagger).$$

\mathcal{H}_1 e \mathcal{H}_2 sono operatori autoaggiunti e commutanti, cui si può applicare il teorema precedente. Indipendentemente dal Teor.3.3.9 si può argomentare anche come segue: sia

$$(\mathcal{N} - \lambda \mathbb{1})^\nu |\xi\rangle = 0$$

Mostriamo che ciò implica $(\mathcal{N} - \lambda \mathbb{1})|\xi\rangle = 0$ e quindi che il polinomio minimale di \mathcal{N} deve avere tutte radici semplici. Sia $\mathcal{N}_\lambda \equiv \mathcal{N} - \lambda \mathbb{1}$ e per ogni k compreso tra 0 e ν troviamo

$$\|\mathcal{N}_\lambda^{\nu-k} \mathcal{N}_\lambda^{\dagger k} |\xi\rangle\| = \langle \xi | \mathcal{N}_\lambda^k \mathcal{N}_\lambda^{\dagger \nu-k} \mathcal{N}_\lambda^{\nu-k} \mathcal{N}_\lambda^{\dagger k} |\xi \rangle = \|\mathcal{N}_\lambda^\nu |\xi\rangle\| = 0.$$

Ne segue $\mathcal{N}_\lambda^{\nu-k} \mathcal{N}_\lambda^{\dagger k} |\xi\rangle = 0$ e anche $(\mathcal{N}_\lambda \pm \mathcal{N}_\lambda^\dagger)^\nu |\xi\rangle = 0$. Ma $(\mathcal{N}_\lambda + \mathcal{N}_\lambda^\dagger)$ e $i(\mathcal{N}_\lambda - \mathcal{N}_\lambda^\dagger)$ sono autoaggiunti per cui si conclude che $\mathcal{N}_\lambda |\xi\rangle = 0$. ■

3.4. Funzioni di più operatori lineari

3.4.1. Insiemi completi di operatori autoaggiunti. Consideriamo un certo numero ℓ di operatori autoaggiunti $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_\ell$ a due a due commutanti. Ad ognuno di essi è associata una decomposizione spettrale

$$\mathcal{A}_j = \sum_i \lambda_i^{(j)} \mathbb{P}_i^{(j)}.$$

LEMMA. I proiettori $\mathbb{P}_i^{(j)}$ commutano tutti a due a due.

PROVA. È noto che i proiettori sono esprimibili attraverso la formula di Riesz–Dunford in termini del risolvante dell'operatore e quindi sono funzioni dell'operatore; la commutatività è quindi diretta conseguenza della commutatività degli operatori \mathcal{A}_j . ■

Teor.3.4.1

TEOREMA 3.4.1. *Gli operatori*

$$\mathbb{P}_{i_1 i_2 \dots i_\ell} \equiv \mathbb{P}_{i_1}^{(1)} \mathbb{P}_{i_2}^{(2)} \dots \mathbb{P}_{i_\ell}^{(\ell)}$$

sono proiettori ortogonali e cioè

$$\mathbb{P}_{i_1 i_2 \dots i_\ell} \mathbb{P}_{j_1 j_2 \dots j_\ell} = \mathbb{P}_{i_1 i_2 \dots i_\ell} \delta_{i_1 j_1} \delta_{i_2 j_2} \dots \delta_{i_\ell j_\ell}$$

(immediata conseguenza del lemma).

Data una funzione di più variabili $f(z_1, z_2, \dots, z_\ell)$ definiamo una funzione degli operatori \mathcal{A}_j come segue

$$\hat{f}(\mathcal{A}_1, \dots, \mathcal{A}_\ell) \equiv \sum_{\{i\}} f(\lambda_{i_1}, \lambda_{i_2}, \dots, \lambda_{i_\ell}) \mathbb{P}_{i_1 i_2 \dots i_\ell}$$

Def.3.4.1

DEF. 3.4.1. Gli operatori $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_\ell$ si dicono costituire un *insieme completo* di operatori autoaggiunti commutanti se ogni proiettore $\mathbb{P}_{i_1 i_2 \dots i_\ell}$ proietta su un sottospazio mono-dimensionale oppure è l'operatore nullo.

Teor.3.4.2

TEOREMA 3.4.2. *Gli operatori $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_\ell$ formano un insieme completo di operatori autoaggiunti commutanti se e solo se ogni altro operatore autoaggiunto \mathcal{B} che commuti con tutti gli \mathcal{A}_j è una funzione di $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_\ell$.*

PROVA. Dalla definizione segue che per ogni scelta di autovalori λ_j esiste al più un autovettore $|\lambda_{i_1}^{(1)} \lambda_{i_2}^{(2)} \dots \lambda_{i_\ell}^{(\ell)}\rangle$; Dalla commutatività segue che anche $\mathcal{B} |\lambda_{i_1}^{(1)} \lambda_{i_2}^{(2)} \dots \lambda_{i_\ell}^{(\ell)}\rangle$ è autovettore di \mathcal{A}_j appartenente agli stessi autovalori e perciò

$$\mathcal{B} |\lambda_{i_1}^{(1)} \lambda_{i_2}^{(2)} \dots \lambda_{i_\ell}^{(\ell)}\rangle = b(\{\lambda\}) |\lambda_{i_1}^{(1)} \lambda_{i_2}^{(2)} \dots \lambda_{i_\ell}^{(\ell)}\rangle.$$

Il fattore numerico $b(\lambda_{i_1}^{(1)} \lambda_{i_2}^{(2)} \dots \lambda_{i_\ell}^{(\ell)})$ definisce \mathcal{B} come funzione

$$\mathcal{B} = \sum_{i_1 i_2 \dots i_\ell} b(\lambda_{i_1}^{(1)} \lambda_{i_2}^{(2)} \dots \lambda_{i_\ell}^{(\ell)}) \mathbb{P}_{i_1 i_2 \dots i_\ell} = b(\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_\ell).$$

■

⚡ È chiaro che se qualche $\mathbb{P}_{i_1 i_2 \dots i_\ell}$ proietta su un sottospazio a dimensione maggiore di uno esisteranno operatori commutanti con tutti gli \mathcal{A}_j , ma non esprimibili in termini degli stessi proiettori. In genere si aggiungono allora uno o più operatori per “completare” l'insieme. La definizione di insieme completo è importante in Meccanica Quantistica (si parla allora di “insiemi completi di osservabili commutanti”). ◀

3.4.2. Insiemi irriducibili di operatori. Siano $\mathcal{B}_1, \dots, \mathcal{B}_\ell$ operatori autoaggiunti in \mathfrak{C} aventi la seguente proprietà: non esiste alcun sottospazio proprio $\mathfrak{C}' \subset \mathfrak{C}$ invariante rispetto a $\mathcal{B}_1, \dots, \mathcal{B}_\ell$. Si dice allora che gli operatori formano un *insieme irriducibile*.

SCHUR

LEMMA DI SCHUR. *Sia \mathcal{E} un operatore che commuti con tutti gli operatori di un insieme irriducibile $\mathcal{B}_1, \dots, \mathcal{B}_\ell$. Allora \mathcal{E} è multiplo dell'identità ($\mathcal{E} \propto \mathbb{1}$).*

PROVA. Sia λ un autovalore di \mathcal{E} . Il sottospazio \mathcal{S}_λ individuato dagli autovettori di \mathcal{E} appartenenti a λ è lasciato invariante dagli operatori \mathcal{B}_k e per l'ipotesi di irriducibilità \mathcal{S}_λ deve coincidere con tutto lo spazio \mathfrak{C} e dunque $\mathcal{E} = \lambda \mathbb{1}$. ■

ESEMPIO 3-3. Le due matrici Hermitiane

$$J_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, J_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

formano un insieme irriducibile; infatti J_2 ha tre autovalori distinti e perciò ogni operatore \mathcal{E} che commuti con J_2 deve essere diagonale; la commutazione con J_1 impone che $\mathcal{E} \propto \mathbb{1}$:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} - \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & b-a & 0 \\ a-b & 0 & c-b \\ 0 & b-c & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies a = b = c.$$

3.4.3. Funzioni di operatori non-commutanti. Abbiamo visto che funzioni di operatori commutanti si possono definire facilmente attraverso la decomposizione spettrale. Il problema si complica, ma diventa più interessante per le applicazioni, se si vuole definire una funzione di operatori non-commutanti. Diamo di seguito alcune formule che entrano spesso nelle applicazioni (dalla Meccanica Quantistica agli algoritmi di Meccanica Celeste), senza pretendere di discutere il problema nella sua generalità.

3.4.3.1. *La formula di Lie.* Siano \mathcal{A} e \mathcal{B} due operatori lineari e $\mathcal{U} = \exp(\mathcal{A} + \mathcal{B})$. La formula di Lie permette di esprimere \mathcal{U} in termini di $\exp(\mathfrak{t}\mathcal{A})$ e $\exp(\mathfrak{t}\mathcal{B})$ anche nel caso in cui \mathcal{A} e \mathcal{B} non commutino.

Teor.LIE

TEOREMA 3.4.3. Per ogni coppia di operatori lineari \mathcal{A} e \mathcal{B} vale la formula

Teor.3.4.3

$$\exp(\mathcal{A} + \mathcal{B}) = \lim_{n \rightarrow \infty} (e^{\mathcal{A}/n} e^{\mathcal{B}/n})^n$$

PROVA. Partiamo dall'identità

$$\mathcal{A}^n - \mathcal{B}^n \equiv \sum_{k=0}^{n-1} \mathcal{A}^k (\mathcal{A} - \mathcal{B}) \mathcal{B}^{n-k-1}.$$

Siano allora $\mathcal{S}_t \equiv \exp(\mathfrak{t}(\mathcal{A} + \mathcal{B}))$, $\mathcal{T}_t \equiv \exp(\mathfrak{t}\mathcal{A}) \exp(\mathfrak{t}\mathcal{B})$; ponendo $\mathfrak{t} = 1/n$ avremo

$$\begin{aligned} \|\mathcal{S}_t^n - \mathcal{T}_t^n\| &= \left\| \sum_{k=0}^{n-1} \mathcal{S}_t^k (\mathcal{S}_t - \mathcal{T}_t) \mathcal{T}_t^{n-k-1} \right\| \\ &\leq \sum_{k=0}^{n-1} \|\mathcal{S}_t\|^k \|\mathcal{S}_t - \mathcal{T}_t\| \|\mathcal{T}_t\|^{n-k-1} \\ &\leq n \|\mathcal{S}_t - \mathcal{T}_t\| \max(\|\mathcal{S}_t\|, \|\mathcal{T}_t\|)^{n-1} \end{aligned}$$

Sia $\|\mathcal{S}_t\|$ che $\|\mathcal{T}_t\|$ sono minori di $\exp(\frac{1}{n}(\|\mathcal{A}\| + \|\mathcal{B}\|))$, e perciò

$$\|\mathcal{S}_t^n - \mathcal{T}_t^n\| \leq n \|\mathcal{S}_t - \mathcal{T}_t\| \exp(\|\mathcal{A}\| + \|\mathcal{B}\|).$$

Si tratta di mostrare che il limite di quest'ultima espressione per $n \rightarrow \infty$ vale 0. Sviluppando in serie si trova

$$\mathcal{S}_t - \mathcal{T}_t = \frac{1}{2n^2} [\mathcal{A}, \mathcal{B}] + \mathcal{O}(1/n^3)$$

e perciò $\lim_{n \rightarrow \infty} (n \|\mathcal{S}_t - \mathcal{T}_t\|) = 0$. ■

3.4.3.2. *Trasformazioni di similitudine.* Per ogni coppia di operatori lineari \mathcal{A} e \mathcal{B} vale lo sviluppo

$$(3.13) \quad \mathcal{F}(t) = e^{t\mathcal{A}} \mathcal{B} e^{-t\mathcal{A}} = \mathcal{B} + t[\mathcal{A}, \mathcal{B}] + \frac{1}{2}t^2[\mathcal{A}, [\mathcal{A}, \mathcal{B}]] + \dots + \frac{t^n}{n!} C_n + \dots$$

dove i termini C_n sono definiti ricorsivamente da

$$C_{n+1} = [\mathcal{A}, C_n], \quad C_0 = \mathcal{B}.$$

PROVA. Differenziando rispetto a t l'identità da dimostrare si ottiene

$$\frac{\partial \mathcal{F}}{\partial t} = \mathcal{A} e^{t\mathcal{A}} \mathcal{B} e^{-t\mathcal{A}} - e^{t\mathcal{A}} \mathcal{B} e^{-t\mathcal{A}} \mathcal{A} = [\mathcal{A}, \mathcal{F}]$$

ma è anche

$$\frac{\partial \mathcal{F}}{\partial t} = \sum_{n=0}^{\infty} \frac{t^n}{n!} C_{n+1}.$$

Eguagliando i coefficienti in t^n si ottiene il risultato cercato. ■

◀ Come corollario si ha che per ogni funzione f per cui esista $\hat{f}(\mathcal{B})$ vale l'identità $e^{\mathcal{A}} \hat{f}(\mathcal{B}) e^{-\mathcal{A}} = \hat{f}(\mathcal{B} + [\mathcal{A}, \mathcal{B}] + \dots)$ ▶

◀ È molto diffusa nella letteratura matematica la notazione

$$\text{ad}_X(Y) \equiv [X, Y]$$

per cui la Eq. (3.13) si può scrivere più elegantemente

$$e^{t\mathcal{A}} \mathcal{B} e^{-t\mathcal{A}} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \text{ad}_{\mathcal{A}}^n(\mathcal{B}) \equiv \exp\{t \text{ad}_{\mathcal{A}}\}(\mathcal{B}).$$

La terminologia discende dalla teoria dei gruppi di Lie; ad_X corrisponde alla rappresentazione aggiunta dell'algebra di Lie. ▶

B.C.H.

3.4.3.3. *Formula di Baker–Campbell–Hausdorff.* Tenendo conto dei risultati precedenti, calcoliamo ora $\exp(\mathcal{A} + \mathcal{B}) \exp(-\mathcal{A}) \exp(-\mathcal{B})$ nel caso particolarmente semplice in cui \mathcal{A} e \mathcal{B} commutino con $[\mathcal{A}, \mathcal{B}]$:

$$\begin{aligned} e^{\mathcal{A}+\mathcal{B}} &= \lim_{n \rightarrow \infty} (\exp(\mathcal{A}/n) \exp(\mathcal{B}/n))^n \\ &= \lim_{n \rightarrow \infty} e^{\mathcal{A}/n} e^{\mathcal{B}/n} e^{\mathcal{A}/n} e^{\mathcal{B}/n} \dots e^{\mathcal{A}/n} e^{\mathcal{B}/n} \\ &= \lim_{n \rightarrow \infty} \prod_{j=1}^n (e^{j\mathcal{A}/n} e^{\mathcal{B}/n} e^{-j\mathcal{A}/n}) e^{\mathcal{A}} \\ &= \lim_{n \rightarrow \infty} \prod_{j=1}^n \exp\left\{\frac{1}{n}\mathcal{B} + \frac{j}{n^2}[\mathcal{A}, \mathcal{B}]\right\} e^{\mathcal{A}} \\ &= \lim_{n \rightarrow \infty} e^{\mathcal{B}} e^{\mathcal{A}} \exp\left\{\sum \frac{j}{n^2}[\mathcal{A}, \mathcal{B}]\right\} = e^{\mathcal{B}} e^{\mathcal{A}} e^{\frac{1}{2}[\mathcal{A}, \mathcal{B}]} \end{aligned}$$

Questa formula viene applicata al caso degli operatori “canonici” che soddisfano le regole di commutazione di Heisenberg $[\mathbf{a}, \mathbf{a}^\dagger] = \mathbb{1}$. Tuttavia gli operatori canonici esistono solo nello spazio di Hilbert (∞ -dimensionale, vedi Cap. 4.3.3). Ci si può chiedere se, almeno, esistano matrici finite cui applicare la formula di BCH: $[\mathcal{A}, [\mathcal{A}, \mathcal{B}]] = [\mathcal{B}, [\mathcal{A}, \mathcal{B}]] = 0$. La

risposta è affermativa tuttavia queste condizioni pongono severe restrizioni sulle matrici \mathcal{A} e \mathcal{B} (vedi Eserc. 73 e segg.).

In generale il prodotto $e^{\mathcal{A}}e^{\mathcal{B}}$ è esprimibile come uno sviluppo infinito nella forma:

$$(3.14) \quad e^{\mathcal{A}}e^{\mathcal{B}} = \exp \{ \mathcal{H}(\mathcal{A}, \mathcal{B}) \}$$

$$\mathcal{H}(\mathcal{A}, \mathcal{B}) = \mathcal{A} + \mathcal{B} + \frac{1}{2}[\mathcal{A}, \mathcal{B}] + \frac{1}{12}[\mathcal{A}, [\mathcal{A}, \mathcal{B}]] + \frac{1}{12}[\mathcal{B}, [\mathcal{B}, \mathcal{A}]] - \frac{1}{24}[\mathcal{A}, [\mathcal{B}, [\mathcal{A}, \mathcal{B}]]] + \dots$$

Per lo sviluppo completo si vedano le note alla fine del capitolo e anche il Probl. 82.

3.4.4. Teoria delle perturbazioni.

3.4.4.1. *Perturbazioni stazionarie - caso non-degenere.* Sia $\mathcal{H} = \mathcal{H}_0 + \varepsilon \mathcal{V}$ un operatore autoaggiunto con \mathcal{H}_0 e \mathcal{V} entrambi autoaggiunti e ε sia un parametro reale. Il calcolo dello spettro di \mathcal{H} si può ridurre a quello di \mathcal{H}_0 se è sufficiente conoscere il risultato in forma approssimata nell'ipotesi che ε sia "sufficientemente piccolo". Ciò è non solo facile da realizzare, ma in taluni casi è l'unico modo di ottenere lo spettro, a parte ricorrere al calcolo numerico. Ad esempio si consideri

$$\mathcal{H} \leftrightarrow \begin{pmatrix} 2+\varepsilon v_1 & -1 & 0 & \dots\dots\dots & 0 \\ -1 & 2+\varepsilon v_2 & -1 & 0 & \dots\dots\dots & 0 \\ 0 & -1 & 2+\varepsilon v_3 & -1 & 0 & \dots\dots\dots & 0 \\ \vdots & \dots\dots\dots & \ddots & \ddots & \ddots & \dots\dots\dots & \vdots \\ 0 & \dots\dots\dots & 0 & -1 & 2+\varepsilon v_{n-2} & -1 & 0 \\ 0 & \dots\dots\dots & 0 & -1 & 2+\varepsilon v_{n-1} & -1 & 0 \\ 0 & \dots\dots\dots & 0 & -1 & 2+\varepsilon v_n & -1 & 0 \end{pmatrix}$$

La matrice ridotta a $\varepsilon = 0$ è diagonalizzabile e se ne conoscono autovalori ed autovettori (vedi il § 1.3). Sia dunque E_0 l'autovalore di cui vogliamo valutare la correzione e $|E_0\rangle$ il corrispondente autovettore. Dobbiamo determinare autostato e autovalore dell'Hamiltoniano completo

$$H|E_\varepsilon\rangle = E_\varepsilon|E_\varepsilon\rangle$$

tali che per $\varepsilon \rightarrow 0$ valgano i due limiti

$$\lim_{\varepsilon \rightarrow 0} E_\varepsilon = E_0, \quad \lim_{\varepsilon \rightarrow 0} |E_\varepsilon\rangle = |E_0\rangle.$$

L'ipotesi che si fa a questo punto è la seguente: *autovalori ed autovettori di H ammettono uno sviluppo in serie di potenze nel parametro ε .* L'ipotesi è del tutto ragionevole trattandosi di soluzioni di equazioni algebriche in cui ε compare come parametro, tuttavia la natura della serie di potenze è a priori puramente formale.

n \neq ∞ Nel caso a infinite dimensioni la convergenza della **serie perturbativa**, come verranno d'ora in poi denominate le serie di potenze in teoria delle perturbazioni, è da verificare caso per caso, e di norma non si avvera. Anche negli esempi più semplici ci si trova di fronte a serie divergenti ma che non di meno forniscono un utilissimo strumento di calcolo approssimato in quanto **serie asintotiche**. A questo scopo è necessario però disporre di un grande numero di coefficienti della serie, il che non è sempre possibile (si vedano [BW69, OD99]). ◀

Nel seguito affronteremo il problema di base, quello cioè di determinare i coefficienti della serie perturbativa. Cominciamo col definire gli sviluppi

$$(3.15) \quad |E_\varepsilon\rangle = |E_0\rangle + \varepsilon |\eta_1\rangle + \varepsilon^2 |\eta_2\rangle + \dots$$

$$(3.16) \quad E_\varepsilon = E_0 + \delta_1 \varepsilon + \delta_2 \varepsilon^2 + \dots$$

dove i coefficienti reali δ_k e i vettori $|\eta_1\rangle, |\eta_2\rangle, \dots$ sono da determinarsi e sostituiamo il tutto nell'equazione agli autovalori. Identificando i coefficienti delle varie potenze in ε si ottiene, almeno in linea di principio, la soluzione. In realtà è preferibile procedere con un sistema di tipo ricorsivo, in base al quale, nota la soluzione all'ordine n , si ricava facilmente l'ordine $n+1$ (è questo d'altronde il modo in cui si imposta di norma il calcolo, dato che per il fisico poco importa conoscere la soluzione a un dato ordine ε^k senza conoscere il contributo in ε^{k-1} , a priori più rilevante). La relazione di ricorrenza si determina in questo modo (vedi [Sak90, OD99]): riordiniamo i termini dell'equazione per ottenere

$$(3.17) \quad (H_0 - E_0) |E_\varepsilon\rangle = (E_\varepsilon - E_0) |E_\varepsilon\rangle - \varepsilon V |E_\varepsilon\rangle .$$

Prendendo il prodotto scalare dello stato $|E_0\rangle$ con ambo i membri si avrà

$$0 = (E_\varepsilon - E_0) \langle E_0 | E_\varepsilon \rangle - \varepsilon \langle E_0 | V | E_\varepsilon \rangle$$

(avendo sfruttato il fatto che H_0 è autoaggiunto) e quindi

$$(3.18) \quad E_\varepsilon - E_0 = \varepsilon \frac{\langle E_0 | V | E_\varepsilon \rangle}{\langle E_0 | E_\varepsilon \rangle} .$$

Convienne a questo punto adottare temporaneamente una normalizzazione diversa dal consueto per l'autovettore $|E_\varepsilon\rangle$

$$\langle E_0 | E_\varepsilon \rangle = 1,$$

una convenzione che semplifica notevolmente lo sviluppo delle formule. A conti fatti sarà ovviamente possibile fissare la normalizzazione consueta moltiplicando per un opportuno fattore; l'unica difficoltà potrebbe provenire nell'eventualità che l'autovettore esatto sia ortogonale a $|E_0\rangle$, ma ciò non potrà verificarsi per ε sufficientemente piccolo. Se inseriamo lo sviluppo in serie (3.15), ed uguagliamo i coefficienti ad ogni ordine in ε^k otteniamo infatti

$$(3.19)$$

$$\delta_k = \langle E_0 | V | \eta_{k-1} \rangle$$

Riprendiamo ora l'Eq. (3.17); ci si chiede se siamo autorizzati ad invertire l'operatore $H_0 - E_0$ per ottenere

$$|E_\varepsilon\rangle = |E_0\rangle + \mathcal{R}(E_0) (E_\varepsilon - E_0 - \varepsilon V) |E_\varepsilon\rangle .$$

L'operatore **risolvente** $\mathcal{R}(z) \equiv (H_0 - z\mathbb{1})^{-1}$ è definito per ogni valore reale o complesso di z , ad eccezione dei punti dello spettro di H_0 . Tuttavia i vettori a cui dobbiamo applicare $\mathcal{R}(E_0)$ sono tutti ortogonali al vettore $|E_0\rangle$ e perciò l'inversione è legittima. Si intenderà che il risolvente sottintende la proiezione sul sottospazio ortogonale al vettore $|E_0\rangle$

$$\mathcal{R}(E_0) = \mathbb{P}_0^\perp (H_0 - E_0)^{-1} \mathbb{P}_0^\perp, \quad \left(\mathbb{P}_0^\perp \equiv \mathbb{1} - \sum_{\substack{E \in \sigma(H_0) \\ E \neq E_0}} |E\rangle \langle E| \right) .$$

(Si parla in questi casi di “inverso generalizzato”). Il calcolo del risolvente è ottenibile semplicemente da

$$\mathcal{R}(E_0) |E'_0\rangle = \frac{|E'_0\rangle}{E'_0 - E_0}$$

$$\mathcal{R}(E_0) |E_0\rangle \equiv 0.$$

Se a questo punto inseriamo gli sviluppi in serie nell'equazione precedente otteniamo la seconda relazione di ricorrenza

$$(3.20) \quad |\eta_k\rangle = R_0 \left(\sum_{j=1}^{k-1} \delta_j |\eta_{k-j}\rangle - V |\eta_{k-1}\rangle \right)$$

Le Eq. (3.19), (3.20) formano un *algoritmo ricorsivo* che in linea di principio può generare i coefficienti della serie perturbativa ad ogni ordine per il caso non-degenere. Il vantaggio di questa impostazione è di permettere una notevole economia di calcolo; inoltre l'algoritmo è facilmente codificabile per un calcolo automatico.

Le formule approssimate al 2° ordine sono date da

$$\delta_1 = \langle E_0 | V | E_0 \rangle$$

$$|\eta_1\rangle = -\mathcal{R}(E_0) V |E_0\rangle$$

$$\delta_2 = -\langle E_0 | V \mathcal{R}(E_0) V | E_0 \rangle = \sum_{E'_0 \neq E_0} \frac{|\langle E_0 | V | E'_0 \rangle|^2}{E_0 - E'_0}$$

essendo E'_0 gli autovalori imperturbati diversi da E_0 .

ESEMPIO 3-4. Sia $H = H_0 + \varepsilon \mathcal{V}$, essendo $(\mathcal{H}_0)_{ij} = j \delta_{ij}$ e $\mathcal{V}_{ij} = \delta_{i,j+1} + \delta_{j,i+1}$. L'esercizio è il più semplice possibile, in quanto siamo già in presenza di \mathcal{H}_0 diagonale. Per $N = 10$ si ha ad es.:

$$\mathcal{H} \leftarrow \begin{pmatrix} 1 & \varepsilon & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \varepsilon & 2 & \varepsilon & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \varepsilon & 3 & \varepsilon & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon & 4 & \varepsilon & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varepsilon & 5 & \varepsilon & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \varepsilon & 6 & \varepsilon & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \varepsilon & 7 & \varepsilon & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & 8 & \varepsilon & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & 9 & \varepsilon \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon & 10 \end{pmatrix}$$

La teoria delle perturbazioni al primo ordine dà $\delta_1^{(n)} = 0$ (elementi diagonali della perturbazione). Al secondo ordine

$$\begin{aligned}\delta_2^{(1)} &= \frac{|\mathcal{V}_{12}|^2}{E_0(1) - E_0(2)} = -1 \\ \delta_2^{(2)} &= \frac{|\mathcal{V}_{21}|^2}{E_0(2) - E_0(1)} + \frac{|\mathcal{V}_{23}|^2}{E_0(2) - E_0(3)} = 0 \\ \delta_2^{(3)} &= \frac{|\mathcal{V}_{32}|^2}{E_0(3) - E_0(2)} + \frac{|\mathcal{V}_{34}|^2}{E_0(3) - E_0(4)} = 0 \\ &\dots \dots \\ \delta_2^{(N)} &= \frac{|\mathcal{V}_{N,N-1}|^2}{E_0(N) - E_0(N-1)} = 1\end{aligned}$$

Dunque lo spettro è dato in via approssimata da $E_1 = 1 - \varepsilon^2 + O(\varepsilon^4)$, $E_2 = 2 + O(\varepsilon^4)$, $E_3 = 3 + O(\varepsilon^4)$, ... (perché non $O(\varepsilon^3)$?) il che significa che per avere una stima degli autovalori di indice maggiore di uno siamo costretti a calcolare le correzioni ad alti ordini. Per questo è consigliabile utilizzare un programma automatico: le regole di ricorrenza Eq. (3.19),(3.20) si possono tradurre facilmente in un codice da eseguire su un computer. Nell'archivio raggiungibile via rete si trova un programma **Mathematica** che implementa l'algoritmo. ◀

🔗 Poche righe di **Mathematica** danno il risultato in pochi decimi di secondo utilizzando direttamente la routine di diagonalizzazione:

```
pert.ex1.nb
H0[n_] := SparseArray[{i_, i_} -> i, {n, n}]
V[n_] := SparseArray[{i_, j_} /; Abs[i-j]==1 -> 1}, {n, n}]
H[n_] = H0[n] + a V[n];
Series[Eigenvalues[H[10]], {a, 0, 6}]
```

che fornisce in output ($\varepsilon \rightleftharpoons a$)

$$\begin{aligned}1 - a^2 + \frac{a^4}{2} - \frac{7a^6}{12} + \frac{127a^8}{144} + O(a^{10}) \\ 2 - \frac{a^4}{2} + \frac{2a^6}{3} - \frac{15a^8}{16} + O(a^{10}) \\ 3 - \frac{a^6}{12} + \frac{a^8}{16} + O(a^{10}) \\ 4 - \frac{a^8}{144} + O(a^{10}) \\ 5 + O(a^{10}) \\ \dots\dots\end{aligned}$$

Si nota che al crescere di n (fino a $n = N/2$) l'autovalore E_n presenta una correzione $-\varepsilon^{2n}/(n!(n-1)!)$. Viene spontaneo chiedersi se ci sia un modo semplice di capire questo fatto (vedi il Probl. 81). ◀

🔗 Il programma **Mathematica PertTh.nb**, incluso nel disco in dotazione del libro, permette di risolvere problemi elementari di teoria delle perturbazioni in modo automatico. È naturale chiedersi il motivo di effettuare calcoli di questo genere “con carta e matita” quando l'uso del

computer può alleviare la fatica. Il consiglio è sempre quello di familiarizzarsi con l'argomento senza l'uso del calcolatore e solo in seguito, con le idee chiare, risparmiarsi la fatica. Ai primi passi è addirittura consigliabile l'uso del regolo calcolatore... ◀

ESEMPIO 3-5. Sia $H = H_0 + \varepsilon \mathcal{V}$, essendo, come si è anticipato a Pag. 59, $(\mathcal{H}_0)_{ij} = N^2(2\delta_{i,j} - \delta_{i,j+1} - \delta_{j,i+1})$ e $\mathcal{V}_{ij} = \frac{1}{N^2} j(N-j)\delta_{ij}$, $\dim(\mathfrak{E}) = N-1$. Lo spettro di \mathcal{H}_0 è noto ($E_n = 2N^2(1 - \cos(\frac{n\pi}{N}))$). La teoria delle perturbazioni al primo ordine dà

$$\delta_1^{(n)} = \frac{2}{N} \sum_{j=1}^{N-1} \frac{j(N-j)}{N^2} \left(\sin \frac{nj\pi}{N} \right)^2.$$

mentre al secondo ordine la formula è meno maneggevole

$$\delta_2^{(n)} = \frac{2}{N^4} \sum_{m \neq n} \frac{\left(\sum_j \frac{j(N-j)}{N^2} \sin \frac{nj\pi}{N} \sin \frac{mj\pi}{N} \right)^2}{\cos \frac{n\pi}{N} - \cos \frac{m\pi}{N}}.$$

L'operatore H_0 costituisce una approssimazione discreta all'Hamiltoniano di una particella confinata in una scatola di lato unitario secondo la meccanica quantistica. \mathcal{V} rappresenta l'energia potenziale. Nel package allegato si trova il codice **matlab** (*pert_ex3-5.m*) che calcola lo spettro numericamente e lo confronta con la formula perturbativa. Si possono anche sfruttare le capacità simboliche di **Mathematica** (*pert_ex5.nb*). Siete incoraggiati a esplorare cosa succede al variare dei parametri (N, ε) . ◀

ESEMPIO 3-6. Sia $\mathcal{H} = \mathcal{H}_0 + \varepsilon \mathcal{V}$, essendo $(\mathcal{H}_0)_{ij} = (j-1)\delta_{ij}$ e $\mathcal{V}_{ij} = \delta_{i,j+1}\sqrt{i} + \delta_{j,i+1}\sqrt{j}$; per $N = 10$ si ha ad es.:

$$\mathcal{H} \leftrightarrow \begin{pmatrix} 0 & \varepsilon\sqrt{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \varepsilon\sqrt{1} & 1 & \varepsilon\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \varepsilon\sqrt{2} & 2 & \varepsilon\sqrt{3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon\sqrt{3} & 3 & \varepsilon\sqrt{4} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varepsilon\sqrt{4} & 4 & \varepsilon\sqrt{5} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \varepsilon\sqrt{5} & 5 & \varepsilon\sqrt{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \varepsilon\sqrt{6} & 6 & \varepsilon\sqrt{7} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon\sqrt{7} & 7 & \varepsilon\sqrt{8} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon\sqrt{8} & 8 & \varepsilon\sqrt{9} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varepsilon\sqrt{9} & 9 \end{pmatrix}$$

La teoria delle perturbazioni al primo ordine dà $\delta_1^{(n)} = 0$ (elementi diagonali della perturbazione). Al secondo ordine

$$\begin{aligned}\delta_2^{(1)} &= \frac{|\mathcal{V}_{12}|^2}{E_0(1) - E_0(2)} = -1 \\ \delta_2^{(2)} &= \frac{|\mathcal{V}_{21}|^2}{E_0(2) - E_0(1)} + \frac{|\mathcal{V}_{23}|^2}{E_0(2) - E_0(3)} = -1 \\ \delta_2^{(3)} &= \frac{|\mathcal{V}_{32}|^2}{E_0(3) - E_0(2)} + \frac{|\mathcal{V}_{34}|^2}{E_0(3) - E_0(4)} = -1 \\ &\dots \dots \\ \delta_2^{(N-1)} &= \frac{|\mathcal{V}_{N,N-1}|^2}{E_0(N) - E_0(N-1)} + \frac{|\mathcal{V}_{N-1,N}|^2}{E_0(3) - E_0(4)} = -1 \\ \delta_2^{(N)} &= \frac{|\mathcal{V}_{N,N-1}|^2}{E_0(N) - E_0(N-1)} = N-1\end{aligned}$$

Dunque lo spettro è dato in via approssimata da $E_n = n - 1 - \varepsilon^2 + O(\varepsilon^4)$ per $n < N$ (perché non $O(\varepsilon^3)$?). Si noti che la somma dei coefficienti perturbativi δ_2 deve annullarsi in quanto è zero la traccia di \mathcal{V} . In realtà il problema è affrontabile più efficacemente per via algebrica. Ciò è suggerito dal fatto che la matrice \mathcal{V} rappresenta, a meno di un fattore $\sqrt{2}$, l'operatore di posizione q in Meccanica Quantistica, rappresentato nella base dell'energia dell'oscillatore armonico, e troncato a dimensione finita N . Si trova che, ponendo $\mathbb{P}_n = |n\rangle\langle n|$, gli operatori \mathcal{H}_0 , \mathcal{V} e $\mathcal{W} = [\mathcal{H}_0, \mathcal{V}]$ soddisfano le relazioni di commutazione

$$[\mathcal{W}, \mathcal{H}_0] = -\mathcal{V}, \quad [\mathcal{V}, \mathcal{W}] = 2(\mathbb{1} - (N-1)\mathbb{P}_N)$$

Consideriamo allora la trasformazione di similitudine (vedi Cap. 3.4.3.2)

$$e^{-\varepsilon\mathcal{W}} \mathcal{H}_0 e^{\varepsilon\mathcal{W}} = \mathcal{H}_0 + \varepsilon\mathcal{V} + \varepsilon^2(\mathbb{1} - (N-1)\mathbb{P}_N)$$

Si conclude che \mathcal{H} è unitariamente equivalente a $\mathcal{H}_0 - \varepsilon^2\mathbb{1}$ più una correzione di rango uno (il proiettore) che si può valutare perturbativamente come nel Probl. 81, e risulta $O(\varepsilon^{2(N-n)})$, cioè rappresenta un “effetto di taglia finita” che scompare nel limite $N \rightarrow \infty$ a n fissato e $\varepsilon < 1$. Cosa succede per $\varepsilon \geq 1$? L'operatore \mathcal{V} diventa dominante; il suo spettro è dato dagli zeri del polinomio di Hermite $H_N(x/\sqrt{2})$; il resto è lasciato come esercizio.

🔗 Poche istruzioni di *Mathematica* aiutano a destreggiarsi nei calcoli algebrici dell'esempio precedente. Si veda `pert_ex6.nb`. ◀

3.4.4.2. Autovalori degeneri. Nel caso l'autovalore imperturbato corrisponda a più di un autostato (\mathbb{P}_λ proietta su un sottospazio a dimensione > 1) le formule del paragrafo precedente devono essere modificate. Già il punto di partenza è da studiare attentamente: la base di autovettori è definita a meno di trasformazioni unitarie all'interno del sottospazio degeneri (come succede in un corpo rigido con simmetria assiale, un caso in cui gli assi principali di inerzia non sono definiti univocamente). Perciò si deve per prima cosa individuare la base corretta nel limite di perturbazione zero. Siano $\{|E_0, \alpha\rangle, \alpha = 1, 2, \dots, r\}$ i vettori di base nel sottospazio degeneri appartenenti all'autovalore E_0 . L'equazione (3.17) diventa perciò

$$(H_0 - E_0) |E_\varepsilon, \alpha\rangle = (E_\varepsilon - E_0) |E_\varepsilon, \alpha\rangle - \varepsilon \mathcal{V} |E_\varepsilon, \alpha\rangle$$

e prendendo il prodotto interno con $\langle E_0, \beta |$ otteniamo

$$\varepsilon \langle E_0, \beta | V | E_0, \alpha \rangle = (\varepsilon \delta_1 + O(\varepsilon^2)) \langle E_0, \beta | E_\varepsilon, \alpha \rangle$$

da cui, passando al limite $\varepsilon \rightarrow 0$ segue

$$\langle E_0, \beta | V | E_0, \alpha \rangle = \delta_1 \langle E_0, \beta | E_0, \alpha \rangle = \delta_1 \delta_{\alpha\beta}.$$

Ciò mostra che la correzione al primo ordine è data dagli autovalori della matrice V ristretta al sottospazio di degenerazione e i relativi autovettori definiscono la corretta base su cui costruire il calcolo perturbativo.

ESEMPIO 3-7. Sia

$$H \leftrightarrow \begin{pmatrix} 0 & \lambda & 0 & 0 & 0 & 0 \\ \lambda & 1 & \sqrt{2}\lambda & 0 & 0 & 0 \\ 0 & \sqrt{2}\lambda & 1 & \sqrt{3}\lambda & 0 & 0 \\ 0 & 0 & \sqrt{3}\lambda & 2 & 2\lambda & 0 \\ 0 & 0 & 0 & 2\lambda & 2 & \sqrt{3}\lambda \\ 0 & 0 & 0 & 0 & \sqrt{3}\lambda & 3 \end{pmatrix}$$

L'autovalore 0 è non degenere, per cui si applicano le formule Eq. (3.19,3.20): $E_1 \approx -\lambda^2$. L'autovalore 1 invece è doppio e quindi bisogna calcolare lo spettro della matrice di perturbazione ristretta al sottospazio degenere, in questo caso

$$\sigma \left\{ \begin{pmatrix} 0 & \sqrt{2}\lambda \\ \sqrt{2}\lambda & 0 \end{pmatrix} \right\} = \pm \sqrt{2}\lambda$$

Si noterà che la differenza sostanziale tra il caso degenere e quello non-degenere consiste nel fatto che elementi di matrice fuori dalla diagonale possono contribuire al primo ordine perturbativo. ◀

ESEMPIO 3-8.

$$\mathcal{H} \leftrightarrow \begin{pmatrix} 0 & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & 1 & 0 & \lambda & 0 & 0 \\ \lambda & 0 & 1 & 0 & \lambda & 0 \\ 0 & \lambda & 0 & 2 & 0 & \lambda \\ 0 & 0 & \lambda & 0 & 2 & \lambda \\ 0 & 0 & 0 & \lambda & \lambda & 3 \end{pmatrix}$$

In questo caso la matrice di perturbazione ristretta al sottospazio di degenerazione è identicamente nulla e quindi non fissa la base in modo univoco. È necessario allora diagonalizzare la matrice

$$\langle E_0, \alpha | \mathcal{V} \mathcal{R}(E_0) \mathcal{V} | E_0, \beta \rangle$$

essendo al solito $\mathcal{R}(E_0) = \mathbb{P}_0^\perp (H_0 - E_0)^{-1} \mathbb{P}_0^\perp$ il risolvente generalizzato. Nel nostro caso si ha

$$\langle E_0 = 1, \alpha | \mathcal{V} \mathcal{R}(E_0) \mathcal{V} | E_0 = 1, \beta \rangle = - \begin{pmatrix} 0 & \lambda^2 \\ \lambda^2 & 0 \end{pmatrix}, \quad \langle E_0 = 2, \alpha | \mathcal{V} \mathcal{R}(E_0) \mathcal{V} | E_0 = 2, \beta \rangle = \begin{pmatrix} 0 & \lambda^2 \\ \lambda^2 & 0 \end{pmatrix},$$

e perciò $E_1 = 1 \pm \lambda^2 + O(\lambda^3)$, $E_2 = 2 \pm \lambda^2 + O(\lambda^3)$. Anche in questo caso lo spettro si può trovare in forma chiusa (provare con **Mathematica**), ma la potenza del metodo perturbativo sta nella sua applicabilità ben oltre i casi semplici, offerti solo a titolo esemplificativo.

Come generalizzazione provare a studiare il caso $(\mathcal{H}_0)_{ij} = (2\delta_{i,j} - \delta_{i,j+1} - \delta_{j,i+1} - \delta_{1N} - \delta_{N1})$ e $\mathcal{V}_{ij} = \varepsilon \sin(2\pi j) \delta_{ij}$. In questo caso il calcolo perturbativo presenta qualche sottigliezza,

in quanto, come nel caso precedente, la matrice $\langle E_0, \beta | V | E_0, \alpha \rangle$ non individua la base di partenza, essendo identicamente nulla. ◀

3.4.5. Soluzione di sistemi lineari. Uno dei problemi più comuni che si presentano nelle applicazioni è quello della soluzione di sistemi di equazioni lineari del tipo

$$(3.21) \quad \frac{d}{dt} \psi(t) = \mathcal{A} \psi(t)$$

dove $\psi(t)$ è un vettore dipendente dal tempo t e \mathcal{A} è un operatore lineare che può essere a sua volta funzione del tempo. Consideriamo per il momento il caso di \mathcal{A} indipendente dal tempo. Allora è immediato verificare che

$$\psi(t) = \exp\{t\mathcal{A}\} \psi(0)$$

è la soluzione del sistema. L'esponenziale dell'operatore si può, almeno in linea di principio, calcolare con i metodi della teoria spettrale: se $\mathcal{A} = \sum_{\lambda} (\lambda \mathbb{P}_{\lambda} + \mathbb{E}_{\lambda})$ allora sarà

$$\psi(t) = \sum_{\lambda} e^{\lambda t} \mathbb{P}_{\lambda} \left(\mathbb{1} + t \mathbb{E}_{\lambda} + \dots + \frac{t^{\nu_{\lambda}}}{\nu_{\lambda}!} \mathbb{E}_{\lambda}^{\nu_{\lambda}} \right) \psi(0)$$

ovvero, per operatori diagonalizzabili, semplicemente

$$\psi(t) = \sum_{\lambda} e^{\lambda t} \mathbb{P}_{\lambda} \psi(0) .$$

Nella pratica si devono affrontare vari problemi per portare in fondo il calcolo. Tutto dipende dalla dimensione della matrice, dalle sue caratteristiche (diagonalizzabilità, posizione e natura dello spettro, quasi-degenerazione degli autovalori, stabilità rispetto a piccole perturbazioni, ad es. gli errori introdotti dalla aritmetica a numero finito di cifre, etc. - si veda ad es. [ML78]). Se l'operatore è dipendente dal tempo la teoria spettrale non è più applicabile, all'infuori del caso particolare in cui tale dipendenza sia banale e cioè

$$\mathcal{A}(t) = \sum_j \lambda_j(t) \mathbb{P}_j$$

($\mathcal{A}(t)$ è funzione di $\mathcal{A}(0)$, dunque con gli stessi proiettori spettrali \mathbb{P}). In questo caso si avrà

$$\psi(t) = \sum_j e^{\int_0^t \lambda_j(\tau) d\tau} \mathbb{P}_j \psi(0) \equiv \exp \left\{ \int_0^t \mathcal{A}(\tau) d\tau \right\} \psi(0) .$$

In generale tuttavia l'equazione precedente non è soluzione della Eq. (3.21): se $\mathcal{A}(t)$ e $\mathcal{A}(t')$ non possono essere diagonalizzati simultaneamente per ogni coppia (t, t') , non vale la formula di derivazione

$$(?) \quad \frac{d}{dt} \exp \left\{ \int_0^t \mathcal{A}(\tau) d\tau \right\} \stackrel{?}{=} \mathcal{A}(t) \exp \left\{ \int_0^t \mathcal{A}(\tau) d\tau \right\} .$$

Si procede allora nel modo seguente: si definisce $\mathcal{U}(t)$ l'operatore lineare che soddisfa l'equazione

$$\frac{d}{dt} \mathcal{U}(t) = \mathcal{A}(t) \mathcal{U}(t), \quad \mathcal{U}(0) = \mathbb{1} ,$$

e integrando entrambi i membri si ottiene

$$\mathcal{U}(t) = \mathbb{1} + \int_0^t \mathcal{A}(\tau) \mathcal{U}(\tau) d\tau .$$

Quest'ultima equazione si può iterare per ottenere

$$\mathcal{U}(t) = \mathbb{1} + \int_0^t \mathcal{A}(\tau) \mathcal{U}(\tau) d\tau + \int_0^t \mathcal{A}(\tau) \int_0^\tau \mathcal{A}(\tau') \mathcal{U}(\tau') d\tau' d\tau$$

e dopo n iterazioni

$$\begin{aligned} \mathcal{U}(t) = & \mathbb{1} + \int_0^t \mathcal{A}(\tau) \mathcal{U}(\tau) d\tau + \int_0^t \mathcal{A}(\tau) \int_0^\tau \mathcal{A}(\tau') d\tau' d\tau + \dots + \\ & \int_0^t \int_0^{\tau_1} \dots \int_0^{\tau_{n-2}} \mathcal{A}(\tau_1) \mathcal{A}(\tau_2) \dots \mathcal{A}(\tau_{n-1}) d\tau_1 d\tau_2 \dots d\tau_{n-1} \\ & \int_0^t \int_0^{\tau_1} \dots \int_0^{\tau_{n-1}} \mathcal{A}(\tau_1) \mathcal{A}(\tau_2) \dots \mathcal{A}(\tau_n) \mathcal{U}(\tau_n) d\tau_1 d\tau_2 \dots d\tau_n . \end{aligned}$$

Al limite per $n \rightarrow \infty$ si ottiene la soluzione

$$(3.22) \quad \mathcal{U}(t) = \mathbb{1} + \sum_{n=1}^{\infty} \int_0^t \int_0^{\tau_1} \dots \int_0^{\tau_{n-1}} \mathcal{A}(\tau_1) \mathcal{A}(\tau_2) \dots \mathcal{A}(\tau_n) d\tau_1 d\tau_2 \dots d\tau_n$$

L'equazione precedente assomiglia all'esponenziale $\exp \int_0^t \mathcal{A}(t') dt'$ con la differenza fondamentale che nel suo sviluppo in serie compaiono prodotti $\mathcal{A}(t_1) \mathcal{A}(t_2) \dots \mathcal{A}(t_n)$ in cui gli argomenti sono strettamente ordinati $t_1 > t_2 > \dots t_n$. Si introduce di solito l'*operatore di ordinamento temporale* \mathcal{T} che agisce su ogni prodotto $\mathcal{A}(t_1) \mathcal{A}(t_2) \dots \mathcal{A}(t_n)$ permutando i fattori in modo che gli argomenti soddisfino la proprietà di ordinamento. Ad es.

$$\mathcal{T}(\mathcal{A}(t) \mathcal{A}(s)) = \begin{cases} \mathcal{A}(t) \mathcal{A}(s) & \text{per } t > s \\ \mathcal{A}(s) \mathcal{A}(t) & \text{per } t < s \end{cases}$$

L'Eq. 3.22 si può allora scrivere formalmente

$$\boxed{\mathcal{U}(t) = \mathcal{T} \left(\exp \int_0^t \mathcal{A}(s) ds \right)}$$

Si tratta ovviamente di una scrittura compatta in tutto e per tutto equivalente alla Eq. 3.22 e la sua utilità consiste quasi esclusivamente nella sua facile memorizzabilità; solo in casi molto speciali questa espressione può essere sfruttata in modo sintetico senza passare attraverso lo sviluppo in serie. L'Eq. 3.22 è nota in meccanica quantistica come “sviluppo di Dyson” (ma è in realtà dovuta a Vito Volterra).

3.4.6. Rappresentazione di interazione. Supponiamo che l'operatore \mathcal{A} nell'equazione (3.21) si possa spezzare nella somma $\mathcal{A} = \mathcal{H}_0 + \mathcal{V}(t)$, dove \mathcal{H}_0 sia indipendente dal tempo e ne sia noto lo sviluppo spettrale. Inoltre si assuma che la “perturbazione” $\mathcal{V}(t)$ sia “piccola” rispetto ad \mathcal{H}_0 (ad es. $\|\mathcal{V}\| \ll \|\mathcal{H}_0\|$). Si può allora impostare la soluzione dell'Eq. (3.21) per via approssimata. Indichiamo con $\mathcal{U}_0(t)$ l'operatore $\exp\{t\mathcal{H}_0\}$ e cerchiamo di determinare la soluzione nella forma

$$\begin{aligned} \frac{d}{dt} \mathcal{U}(t) &= (\mathcal{H}_0 + \mathcal{V}(t)) \mathcal{U}(t) \\ \mathcal{U}(t) &= \mathcal{U}_0(t) \mathcal{W}(t) \end{aligned}$$

Derivando rispetto a t si ottiene

$$\frac{d}{dt} (\mathcal{U}_0 \mathcal{W}(t)) = \mathcal{H}_0 \mathcal{U}_0 \mathcal{W}(t) + \mathcal{U}_0 \frac{d}{dt} \mathcal{W}(t) = (\mathcal{H}_0 + \mathcal{V}(t)) \mathcal{U}_0(t) \mathcal{W}(t)$$

ossia

$$\frac{d}{dt} \mathcal{W}(t) = (\mathcal{U}_0^{-1} \mathcal{V}(t) \mathcal{U}_0(t)) \mathcal{W}(t)$$

L'equazione è della stessa forma studiata nella sezione precedente e la soluzione sarà data da

$$\mathcal{W}(t) = \mathcal{T} \left(\exp \int_0^t e^{-t'H_0} \mathcal{V}(t') e^{t'H_0} dt' \right)$$

La soluzione dell'equazione è pertanto

$$\psi(t) = e^{tH_0} \mathcal{T} \left(\exp \int_0^t e^{-t'H_0} \mathcal{V}(t') e^{t'H_0} dt' \right) \psi(0).$$

🔗 Come applicazione determiniamo la derivata $\frac{d}{dt} \exp \mathcal{A}(t)$: si ha

$$\frac{d}{ds} \exp \mathcal{A}(s) = \lim_{h \rightarrow 0} \frac{1}{h} (e^{\mathcal{A}(s+h)} - e^{\mathcal{A}(s)})$$

Identificando $\mathcal{A}(s)$ con H_0 e $h \frac{d}{ds} \mathcal{A}(s)$ con \mathcal{V} possiamo applicare lo sviluppo di Dyson:

$$e^{\mathcal{A}(s+h)} = e^{\mathcal{A}(s)} + h \int_0^1 e^{(1-\tau)\mathcal{A}(s)} \frac{d\mathcal{A}(s)}{ds} e^{\tau\mathcal{A}(s)} d\tau + \mathcal{O}(h^2),$$

e perciò

$$\begin{aligned} \frac{d}{ds} \exp \mathcal{A}(s) &= e^{\mathcal{A}(s)} \int_0^1 e^{-\tau\mathcal{A}(s)} \frac{d\mathcal{A}(s)}{ds} e^{\tau\mathcal{A}(s)} d\tau \\ &= e^{\mathcal{A}(s)} \frac{d\mathcal{A}(s)}{ds} - \frac{1}{2} e^{\mathcal{A}(s)} \left[\mathcal{A}(s), \frac{d\mathcal{A}(s)}{ds} \right] + \dots \end{aligned}$$



Appendice II

3.5. Polinomi Ortogonali

3.5.1. Definizioni. Sia \mathcal{P}_N lo spazio lineare costituito dai polinomi in una variabile reale x di grado[†] non superiore a N . Introduciamo un prodotto interno in \mathcal{P}_N come segue

$$\langle p_1 | p_2 \rangle = \int_a^b w(x) \overline{p_1(x)} p_2(x) dx$$

dove $w(x)$ è una funzione positiva quasi dappertutto nell'intervallo (a, b) , che chiameremo “peso” (tale intervallo potrà essere finito o infinito, potendosi sempre ricondurre alle scelte $(-1, 1)$, $(0, +\infty)$, $(-\infty, +\infty)$). I monomi $1, x, x^2, \dots, x^N$ formano una base in \mathcal{P}_N . La matrice metrica è data in questa base da

$$h_{ij} = \int_a^b w(x) x^{i+j} dx \equiv m_{i+j}$$

dove $m_k = \int_a^b w(x) x^k dx$ definiscono i *momenti* della distribuzione $w(x) dx$. Possiamo sempre assumere che la distribuzione sia normalizzata, ossia $m_0 = \int_a^b w(x) dx = 1$. A partire dalla matrice metrica si può costruire una base ortonormale con il procedimento di Schmidt introdotto nella Sez. 2.3. Esplicitamente si ha

$$P_k(x) = \mathcal{N}_k \begin{vmatrix} m_0 & m_1 & m_2 & \dots & m_k \\ m_1 & m_2 & m_3 & \dots & m_{k+1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ m_{k-1} & m_k & m_{k+1} & \dots & m_{2k-1} \\ 1 & x & x^2 & \dots & x^k \end{vmatrix}$$

È facile verificare attraverso un calcolo esplicito che $P_k(x)$ è ortogonale a tutti i polinomi di grado inferiore: sviluppando P_k secondo la formula di Lagrange applicata all'ultima riga,

$$\langle x^j | P_k \rangle = \sum_{i=0}^k G_i^{(k)} \langle x^j | x^i \rangle = \sum_i m_{i+j} G_i^{(k)}$$

avendo indicato con $G_i^{(k)}$ il complemento algebrico di x^i nel determinante che definisce P_k ; indicheremo anche semplicemente con G_n il complemento algebrico della potenza

[†]Il grado massimo è arbitrario e infatti in tutto lo sviluppo della teoria può essere preso grande a piacere e anche infinito, ma questo ci porterebbe direttamente al capitolo seguente.

massima, ossia $G_n^{(n)}$. Ne segue

$$\langle x^j | P_k \rangle = \begin{vmatrix} m_0 & m_1 & m_2 & \dots & m_k \\ m_1 & m_2 & m_3 & \dots & m_{k+1} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ m_{k-1} & m_k & m_{k+1} & \dots & m_{2k-1} \\ m_j & m_{j+1} & m_{j+2} & \dots & m_{j+k} \end{vmatrix}$$

È chiaro che per $0 \leq j < k$ il determinante si annulla avendo due righe uguali. La costante \mathcal{N}_k si può fissare per imporre la normalizzazione di P_k in base a qualche convenzione. Se richiediamo $\|P_k\| = 1$ allora si dimostra che

$$\mathcal{N}_n = (G_n G_{n+1})^{-1/2}, \quad G_n \equiv G_n^{(n)}$$

ESEMPIO 3-9. I polinomi di Legendre sono definiti per $(a, b) = (-1, 1)$, $w(x) = 1/2$.

$$m_k = \frac{1}{2} \int_{-1}^1 x^k dx = \frac{1+(-1)^k}{2(k+1)}$$

$$\|h_{ij}\| = \begin{pmatrix} 1 & 0 & 1/3 & 0 & 1/5 & 0 & 1/7 & 0 & \dots \\ 0 & 1/3 & 0 & 1/5 & 0 & 1/7 & 0 & 1/9 & \dots \\ 1/3 & 0 & 1/5 & 0 & 1/7 & 0 & 1/9 & 0 & \dots \\ 0 & 1/5 & 0 & 1/7 & 0 & 1/9 & 0 & 1/11 & \dots \\ 1/5 & 0 & 1/7 & 0 & 1/9 & 0 & 1/11 & \ddots & \dots \\ 0 & 1/7 & 0 & 1/9 & 0 & 1/11 & 0 & \ddots & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \dots \end{pmatrix}$$

I primi polinomi (di grado più basso) si calcolano rapidamente:

$$P_1(x) = 1$$


$$P_2(x) = \begin{vmatrix} 1 & 0 \\ 1 & x \end{vmatrix} = x$$

$$P_3(x) = \begin{vmatrix} 1 & 0 & 1/3 \\ 0 & 1/3 & 0 \\ 1 & x & x^2 \end{vmatrix} = \frac{1}{3} \left(x^2 - \frac{1}{3} \right)$$

$$P_4(x) = \begin{vmatrix} 1 & 0 & 1/3 & 0 \\ 0 & 1/3 & 0 & 1/5 \\ 1/3 & 0 & 1/5 & 0 \\ 1 & x & x^2 & x^3 \end{vmatrix} = \frac{4}{135} \left(x^3 - \frac{3}{5} x \right)$$

Per ottenere una base ortonormale si tratta ora di moltiplicare ciascun polinomio per la costante di normalizzazione \mathcal{N}_n . Otteniamo così

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= \sqrt{3}x \\ P_2(x) &= \frac{1}{2}\sqrt{5}(3x^2 - 1) \\ P_3(x) &= \frac{1}{2}\sqrt{7}(5x^3 - 3x) \\ P_4(x) &= \frac{1}{8}\sqrt{9}(35x^4 - 30x^2 + 3) \\ &\dots \end{aligned}$$

Vedremo tuttavia che esistono tecniche più efficienti per calcolare i polinomi, ad es. quella delle relazioni di ricorrenza e delle funzioni generatrici. 

3.5.2. Relazioni di ricorrenza.

TEOREMA 3.5.1. *I polinomi ortogonali $\{P_n(x)\}$ definiti rispetto a qualunque peso $w(x)$ soddisfano ad una relazione di ricorrenza del tipo*

Teor. 3.5.1

$$(3.23) \quad x P_j(x) = a_j P_{j+1}(x) + b_j P_j(x) + c_j P_{j-1}(x)$$

PROVA. I polinomi $P_n(x)$ formano una base, pertanto dovrà esistere uno sviluppo del tipo

$$x P_n(x) = \sum_{j=0}^{n+1} \alpha_{jn} P_j(x).$$

I coefficienti dello sviluppo si possono ottenere prendendo il prodotto interno con $P_j(x)$:

$$\langle P_j | x P_n \rangle = \alpha_{jn} \|P_j\|^2$$

Tuttavia si ha $\langle P_j | x P_n \rangle = \langle x P_j | P_n \rangle$, come si controlla facilmente tenendo conto che i polinomi sono a valori reali. Ne segue che

$$\alpha_{jn} = 0 \quad \text{per } j+1 < n$$

e quindi lo sviluppo contiene solo tre termini. 

Supponiamo ora che i polinomi siano normalizzati, cioè $\|P_n\| = 1$, e indichiamo con $\kappa_n = \sqrt{G_n/G_{n+1}}$ il coefficiente di x^n di $P_n(x)$. Si ha allora

$$x P_j(x) = x (\kappa_j x^j + \dots) = \frac{\kappa_j}{\kappa_{j+1}} (\kappa_{j+1} x^{j+1} + \dots)$$

il che mostra che nella Eq. (3.23) si deve avere $a_j = \frac{\kappa_j}{\kappa_{j+1}}$. Di più, tenendo conto della relazione di simmetria si ha

$$c_j = \langle P_{j-1} | x P_j \rangle = \langle x P_{j-1} | P_j \rangle = \frac{\kappa_{j-1}}{\kappa_j}.$$

I coefficienti b_k sono calcolabili, non così semplicemente, a partire dal coefficiente di x^{n-1} [Erd55]; nel caso in cui il peso $w(x)$ e l'intervallo (a, b) siano simmetrici rispetto a $x = 0$ si ha $b_k = 0$ in quanto i polinomi sono *a parità definita* (ossia P_n contiene solo potenze pari di x se n è pari, e solo potenze dispari se n è dispari).

PROBLEMA 3-1. *Determinare la forma dei polinomi ortogonali nel caso $(a, b) = (-1, 1)$, $w(x) = \sqrt{1-x^2}$.*

PROBLEMA 3-2. *Determinare la forma dei polinomi ortogonali nel caso $(a, b) = (-1, 1)$, $w(x) = 1/\sqrt{1-x^2}$.*

PROBLEMA 3-3. *Determinare la forma dei polinomi ortogonali nel caso $(a, b) = (-1, 1)$, $w(x) = (1-x)^{\alpha-1} (1+x)^{\beta-1}$.*

3.5.3. Zeri dei polinomi ortogonali.

Teor. 3.5.2

TEOREMA 3.5.2. *Ogni polinomio $P_n(x)$ di una base ortogonale ha tutti i suoi zeri reali e semplici*

PROVA. Supponiamo che P_n abbia m zeri reali di molteplicità dispari, indicati con x_1, \dots, x_m , più altri eventuali zeri di molteplicità pari. Attraversando ciascuno di questi punti P_n cambia segno; ora sia $Q(x) = (x-x_1)(x-x_2)\dots(x-x_m)$. Dato che Q cambia segno in ciascun punto x_j , il prodotto $Q P_n$ ha sempre lo stesso segno. Dunque il prodotto scalare $\langle P_n | Q \rangle$ non può annullarsi, ma allora Q deve avere un grado almeno pari a quello di P_n , in caso contrario dovrebbe essere ortogonale a P_n . Ne segue che P_n ha tutti zeri semplici e reali. ■

Esiste un metodo molto efficace per determinare gli zeri dei polinomi di una qualunque famiglia ortogonale dovuto a Golub e Welsch [GW69]; si riduce il problema a quello di determinare lo spettro di una matrice tridiagonale simmetrica, il che può essere affrontato con metodi molto efficienti (vedi [GL96]).

Partiamo dalla relazione di ricorrenza

$$x P_n(x) = a_n P_{n+1}(x) + b_n P_n(x) + c_n P_{n-1}(x)$$

Fissiamo un n massimo, diciamo N . Allora la relazione si può riscrivere nel formato seguente

$$\begin{pmatrix} b_0 & a_0 & & & & \\ c_1 & b_1 & a_1 & & & \\ \dots & c_2 & b_2 & a_2 & & \\ \dots & & \ddots & \ddots & \ddots & \\ \dots & & & c_k & b_k & a_k \\ \dots & & & & \ddots & \ddots \\ \dots & & & & & c_{n-1} & b_{n-1} & a_{n-1} \\ \dots & & & & & & c_n & b_n \end{pmatrix} \begin{pmatrix} P_0(x) \\ P_1(x) \\ P_2(x) \\ \vdots \\ P_k(x) \\ \vdots \\ P_{n-1}(x) \\ P_n(x) \end{pmatrix} = x \begin{pmatrix} P_0(x) \\ P_1(x) \\ P_2(x) \\ \vdots \\ P_k(x) \\ \vdots \\ P_{n-1}(x) \\ P_n(x) \end{pmatrix} - a_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ P_{n+1}(x) \end{pmatrix}$$

Questa relazione parla da sola! Ogni zero x_k di $P_{n+1}(x)$ è autovalore della matrice a primo membro costruita con i coefficienti della relazione di ricorrenza. Gli zeri si possono dunque determinare diagonalizzando questa matrice, e per questo scopo sono disponibili algoritmi molto efficienti. Il fatto poi che gli zeri siano semplici fa sì che la accuratezza numerica sia ottima. Nei programmi a disposizione su rete si troveranno i codici `matlab` per calcolare gli zeri e i pesi di integrazione Gaussiana, definiti nel seguito, per tutti i polinomi “classici”. Di passaggio notiamo che il fatto che gli zeri siano tutti distinti concorda con il carattere della matrice: gli autovalori di una matrice tridiagonale con elementi tutti non nulli fuori della diagonale principale non possono essere degeneri.

🔗 L'idea che permette di calcolare gli zeri dei polinomi ortogonali si può utilizzare per altre famiglie di funzioni speciali che soddisfanno a una relazione di ricorrenza a tre termini, ad es. alle funzioni di Bessel (vedi il Probl.36, a Pag. 154). ◀

3.5.4. Formule di quadratura di Gauss. La teoria dei polinomi ortogonali permette di realizzare formule di integrazione molto efficienti. Il risultato è dovuto a Gauss. Sia $P_n(x)$ un polinomio di una base ortonormale nell'intervallo (a, b) con peso $w(x)$. Sia $F(x)$ una funzione definita in (a, b) continua e derivabile un numero qualunque di volte. Ricaveremo ora una formula approssimata per l'integrale

$$\int_a^b F(x) w(x) dx$$

basata sulle proprietà dei polinomi ortogonali (seguiamo direttamente la trattazione di [Hoc71]). Cominciamo a considerare il caso in cui F sia un polinomio di grado $2n-1$ e definiamo allora

$$\varphi(x) = \sum_{j=1}^n F(x_j) \frac{P_n(x)}{(x-x_j)P'_n(x)}$$

avendo indicato con x_1, \dots, x_n gli n zeri reali e distinti di $P_n(x)$ e $P'_n(x) = dP_n/dx$. Per il teorema dell'Hôpital si ha che F e φ coincidono in tutti i punti x_j ; allora la differenza $F - \varphi$ è divisibile per $P_n(x)$

$$F(x) = \varphi(x) + P_n(x) Q(x)$$

dove Q è un polinomio di grado $n-1$. Integrando ora entrambi i membri si ha

$$\begin{aligned} \int_a^b F(x) w(x) dx &= \int_a^b w(x) \sum_j F(x_j) \frac{P_n(x)}{(x-x_j)P'_n(x)} dx \\ &\quad + \int_a^b P_n(x) Q(x) w(x) dx \end{aligned}$$

ma Q è ortogonale a P_n e quindi finalmente

$$= \sum_j F(x_j) \frac{1}{P'_n(x)} \int_a^b \frac{w(x)}{x-x_j} P_n(x) dx$$

I coefficienti

$$\lambda_{n,j} = \frac{1}{P'_n(x)} \int_a^b \frac{w(x)}{x-x_j} P_n(x) dx$$

si possono calcolare una volta per tutte essendo indipendenti da $F(x)$. Si chiamano *pesi della formula di quadratura di Gauss*. Si ha in definitiva l'identità

$$\int_a^b F(x) w(x) dx = \sum_j \lambda_{n,j} F(x_j)$$

che rappresenta una formula di quadratura esatta per tutti i polinomi di grado $\leq 2n - 1$. Se applichiamo la formula alle potenze $1, x, x^2, \dots, x^n$ otteniamo il sistema

$$\begin{aligned} 1 &= \sum_j \lambda_{n,j} \\ m_1 &= \sum_j \lambda_{n,j} x_j \\ m_2 &= \sum_j \lambda_{n,j} x_j^2 \\ &\dots \\ m_n &= \sum_j \lambda_{n,j} x_j^n \end{aligned}$$

che ci fornisce il valore dei pesi λ .

ESEMPIO 3-10. Sia $n = 3$ e consideriamo la formula di quadratura Gaussiana per

$$\frac{1}{2} \int_{-1}^1 F(x) dx$$

Gli zeri di P_3 sono $(-\sqrt{3/5}, 0, \sqrt{3/5})$. I pesi si trovano risolvendo il sistema

$$\begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ x_1^2 & x_2^2 & x_3^2 \end{pmatrix} \begin{pmatrix} \lambda_{3,1} \\ \lambda_{3,2} \\ \lambda_{3,3} \end{pmatrix} = \begin{pmatrix} m_0 \\ m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1/3 \end{pmatrix}.$$

Il calcolo offre $\lambda_{3,1} = \lambda_{3,3} = 5/18, \lambda_{3,2} = 4/9$, da cui

$$\int_{-1}^1 F(x) dx = \frac{5}{9} F(-\sqrt{3/5}) + \frac{8}{9} F(0) + \frac{5}{9} F(\sqrt{3/5})$$

formula valida per ogni polinomio di grado ≤ 5 . Se applichiamo la formula, indebitamente, a una generica funzione continua cosa succede? Ad es.

$$\int_{-1}^1 \frac{dx}{3+x} = \ln 2 = .693147...$$



mentre

$$\frac{5}{9} \times \frac{1}{3 + \sqrt{3/5}} + \frac{8}{9} \times \frac{1}{3} + \frac{5}{9} \times \frac{1}{3 - \sqrt{3/5}} \approx .693122...$$



L'esempio precedente suggerisce che se “forziamo” la formula di quadratura inserendo una $F(x)$ che non sia un polinomio di grado $\leq 2n - 1$ si commette un errore, ma che può risultare piccolo. Si dimostra che l'errore è dato da

$$R_n(F) = \left| \int_a^b F(x) w(x) dx - \sum_j F(x_j) \lambda_{n,j} \right| \leq \frac{1}{(2n)!} \max_{a \leq x \leq b} |F^{(2n)}(x)| \kappa_n^2$$

 I codici *Legendre-Gauss.nb* e *Hermite-Gauss.nb* contengono semplici istruzioni *Mathematica* per calcolare i pesi. Questo ha solo uno scopo pedagogico, in quanto *Mathematica* contiene tutta una suite di routine di integrazione anche multidimensionale. 

Mediante la tecnica illustrata nella sezione precedente con cui si calcolano gli zeri, è possibile anche calcolare i pesi λ_j . Si veda ad es. [QSS07]; nel package annesso a queste lezioni si

troveranno le routines `matlab GaussLeg.m`, `GaussHerm.m`,... con le quali si possono calcolare le formula di quadratura di Gauss per tutti i “polinomi classici” che verranno descritti sommariamente nella sezione seguente. Le routines di quadratura di Gauss sono disponibili in molte librerie scientifiche quali la GSL e costituiscono un ingrediente di programmi simbolici quali `Mathematica`. Nell’ottica di non utilizzare programmi a scatola chiusa, lo studente è incoraggiato ad approfondire la logica dei programmi di quadratura gaussiana e di sperimentarne l’affidabilità. Per una trattazione esauriente si vedano [Erd55, Hoc71, Sze59, NO76]. In particolare, è difficile fare meglio di [Erd55], Sec.10.1 cui rimandiamo per documentarsi su tutto quanto riguarda la teoria dei polinomi ortogonali. Quello che segue è un esempio di codice `matlab` per il calcolo delle formule di quadrature di Gauss ($\int_{-\infty}^{\infty} e^{-x^2/2} \dots$):

```

                                GaussHerm.m
function [x,w,p]=gausherm(n)
% Hermite polynomials: zeroes and gaussian integration coefficients.
% Usage:
%       [x,w,p]=gausherm(n)
%
%       returns p=H_j(x_k), j,k=0:n-1 (orthonormal)
%       x roots of H_{n}
%       w Gauss integration weights for the measure
off=sqrt(1:n-1)/sqrt(2);          % off-diag matrix elements
D=diag(off,-1)+diag(off,1);
[U,X]=eig(D);
[x,ind]=sort(-diag(X)');          % zeri in ordine crescente
U=U(:,ind);
p0=U;
w=p0(1,:).^2;                    % Gauss weights
p=U/diag(p0(1,:));               % normalization

```

3.5.5. Polinomi classici. Per i cosiddetti “polinomi classici” (Legendre, Hermite, Laguerre,...) vale una formula esplicita (detta *formula di Rodriguez*):

$$p_n(x) = \frac{1}{w(x)} \left(\frac{d}{dx} \right)^n [w(x) \sigma(x)^n]$$

dove w è il peso e σ è un polinomio. Dalla formula si può ottenere la funzione generatrice della famiglia di polinomi definita da

$$\Phi(x, t) = \sum_{n=0}^{\infty} \frac{1}{n!} p_n(x) t^n$$

che è di aiuto per ricavare molte proprietà dei polinomi. Si parte dalla formula di inversione di Lagrange

$$(3.24) \quad y = x + t \sigma(y) \Rightarrow y(x) = x + \sum_{n=1}^{\infty} \frac{t^n}{n!} \left(\frac{d}{dx} \right)^{n-1} [\sigma(x)^n]$$

o meglio dalla sua generalizzazione

$$f(y) = f(x) + \sum_{n=1}^{\infty} \frac{t^n}{n!} \left(\frac{d}{dx} \right)^{n-1} [f'(x) \sigma(x)^n].$$

Derivando e dopo semplici manipolazioni si ottiene

$$\frac{f'(y) dy}{f'(x) dx} = \frac{1}{f'(x)} \sum_{n=0}^{\infty} \frac{t^n}{n!} \left(\frac{d}{dx} \right)^n [f'(x) \sigma(x)^n]$$

Identificando $f(x)$ con $\int w(x)$ si ottiene infine

$$\frac{w(y)}{w(x)} \frac{dy}{dx} = \frac{1}{w(x)} \sum_{n=0}^{\infty} \frac{t^n}{n!} \left(\frac{d}{dx} \right)^n [w(x) \sigma(x)^n] = \Phi(x, t).$$

Per applicare la formula si deve esprimere y in funzione di x e t dalla Eq. (3.24)

(a) Polinomi di Hermite:

$$w(x) = e^{-x^2}, \sigma(x) = -1, y = x - t, \sum_n \frac{H_n(x)}{n!} t^n = \exp\{-(x-t)^2 + x^2\} = \dots$$

(b) Polinomi di Legendre:

$$w = 1/2, \sigma = \frac{1}{2}(x^2 - 1), y = x - t(y^2 - 1)/2, \Phi(x, t) = (1 - 2xt + t^2)^{-1/2}.$$

(c) Polinomi di Laguerre:

$$w(x) = x^\alpha e^{-x}, \sigma = x, y = x - ty, \Phi(x, t) = (1-t)^{-1-\alpha} \exp\{xt/(t-1)\}.$$

(d) Polinomi di Tchebychev di primo tipo:

$$T_n(x) = \cos(n \cos^{-1} x)$$

$w(x) = 1/\sqrt{1-x^2}$. Sono funzioni trigonometriche sotto altro aspetto, infatti

$$\int_{-1}^1 (1-x^2)^{1/2} T_n(x) T_m(x) dx = \int_0^\pi \cos(n\vartheta) \cos(m\vartheta) = \pi/2 \delta_{nm}$$

(e) Polinomi di Tchebychev di secondo tipo:

$$U_n(x) = \frac{\sin(n+1) \cos^{-1} x}{\sin \cos^{-1} x}$$

$w(x) = \sqrt{1-x^2}$. Sono funzioni trigonometriche sotto altro aspetto, infatti

$$\int_{-1}^1 \sqrt{1-x^2} U_n(x) U_m(x) dx = \int_0^\pi \sin((n+1)\vartheta) \sin((m+1)\vartheta) = \pi/2 \delta_{nm}$$

Si noti che la scelta del polinomio $\sigma(x)$ è obbligata, a meno di banali riscalamanti, per il vincolo che p_n deve essere un polinomio di grado n . Per maggiori dettagli si veda [NO76, Erd55].

AVVERTENZA. La definizione dei polinomi classici offerta qui non è necessariamente allineata con gli standard, ed è offerta solo a scopo introduttivo. Per utilizzare i polinomi in un lavoro serio affidarsi alle sorgenti autorevoli quali [Erd55, GR65, Sze59]. Anche tra un testo e l'altro esistono differenze nella standardizzazione, perciò ... cautela!

3.6. Disuguaglianze fondamentali

Sia \mathcal{C} uno spazio lineare complesso di dimensione n e siano u_1, u_2, \dots, u_n n vettori in \mathcal{C} : si consideri il determinante

$$G(u_1, \dots, u_n) = \det \begin{vmatrix} \langle u_1 | u_1 \rangle & \langle u_1 | u_2 \rangle & \dots & \langle u_1 | u_n \rangle \\ \langle u_2 | u_1 \rangle & \langle u_2 | u_2 \rangle & \dots & \langle u_2 | u_n \rangle \\ \langle u_3 | u_1 \rangle & \langle u_3 | u_2 \rangle & \dots & \langle u_3 | u_n \rangle \\ \dots & \dots & \dots & \dots \\ \langle u_n | u_1 \rangle & \langle u_n | u_2 \rangle & \dots & \langle u_n | u_n \rangle \end{vmatrix}$$

detto *determinante di Gram*.

Teor. 3.6.1 TEOREMA 3.6.1. *Il determinante di Gram è positivo o nullo, inoltre $G = 0$ se e solo se i vettori u_1, u_2, \dots, u_n sono linearmente dipendenti[‡].*

[‡]Nel caso reale, \sqrt{G} coincide con il volume del parallelepipedo avente spigoli u_1, \dots, u_n .

PROVA. Utilizziamo una base ortonormale $\{\mathbf{e}_i\}$. Si ha

$$G_{ij} = \langle \mathbf{u}_i | \mathbf{u}_j \rangle = \sum_k \langle \mathbf{u}_i | \mathbf{e}_k \rangle \langle \mathbf{e}_k | \mathbf{u}_j \rangle.$$

Sia $V_{ik} = \langle \mathbf{u}_i | \mathbf{e}_k \rangle$, allora

$$G_{ij} = \sum_k V_{ik} V_{kj} = (V V^\dagger)_{ij}$$

da cui

$$G = \det(V V^\dagger) \geq 0$$

essendo $G = 0$ solo se V ha determinante nullo, nel qual caso i vettori $\{\mathbf{u}_i\}$ sono linearmente dipendenti. ■

✎ Si noterà che il determinante G_n introdotto a Pag. 70 rappresenta il Gramiano dei vettori di base $1, x, x^2, \dots, x^n$ e pertanto $G_n > 0$. ◀

La disuguaglianza $G \geq 0$ vale anche se il numero di vettori è inferiore alla dimensione dello spazio, in quanto il teorema si applica al sottospazio che contiene i vettori.

DISUGUAGLIANZA DI CAUCHY–SCHWARTZ. Nel caso particolarmente semplice $n = 2$ si trova

$$0 \leq \begin{vmatrix} \langle \mathbf{u}_1 | \mathbf{u}_1 \rangle & \langle \mathbf{u}_1 | \mathbf{u}_2 \rangle \\ \langle \mathbf{u}_2 | \mathbf{u}_1 \rangle & \langle \mathbf{u}_2 | \mathbf{u}_2 \rangle \end{vmatrix} = \|\mathbf{u}_1\|^2 \|\mathbf{u}_2\|^2 - |\langle \mathbf{u}_1 | \mathbf{u}_2 \rangle|^2$$

Cauchy

ossia

$$|\langle \mathbf{u}_1 | \mathbf{u}_2 \rangle| \leq \|\mathbf{u}_1\| \|\mathbf{u}_2\|$$

che è nota come disuguaglianza di Cauchy–Schwartz.

DISUGUAGLIANZA TRIANGOLARE. Segue immediatamente dalla disuguaglianza di Cauchy–Schwartz la disuguaglianza triangolare

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}_1\| + \|\mathbf{u}_2\|.$$

△

DISUGUAGLIANZA DI MINKOWSKI. Per ogni n -pla di numeri complessi (z_1, \dots, z_n) si definisce la p -norma

$$\|z\|_p = \left(\sum_i |z_i|^p \right)^{1/p}$$

essendo p un numero reale ≥ 1 . Il caso $p = 2$ è l'unico in cui la norma corrisponde a un prodotto interno, ma avremo occasione di considerare spesso anche il caso $p = 1$. Formalmente si considera anche il caso $p = \infty$ che per convenzione corrisponde a $\|z\|_\infty = \sup_j |z_j|$.

Vale la disuguaglianza di Minkowski[§]

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p$$

Minkowski

che generalizza la disuguaglianza triangolare.

La dimostrazione farà uso della disuguaglianza di Holder

DISUGUAGLIANZA DI HÖLDER. Siano p e q reali positivi con $\frac{1}{p} + \frac{1}{q} = 1$; allora

$$\sum_i |x_i y_i| \leq \|x\|_p \|y\|_q$$

Holder

[§]http://en.wikipedia.org/wiki/Minkowski_inequality

PROVA. Per ogni coppia di reali positivi a, b si ha

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Ciò discende da $a^p/p = \int_0^a x^{p-1} dx$ e analogo per b . Stante la relazione $\frac{1}{p} + \frac{1}{q} = 1$, la funzione $y = x^{p-1}$ ha come funzione inversa $x = y^{q-1}$ e quindi a^p/p e b^q/q rappresentano le due aree S_1 e S_2 in figura

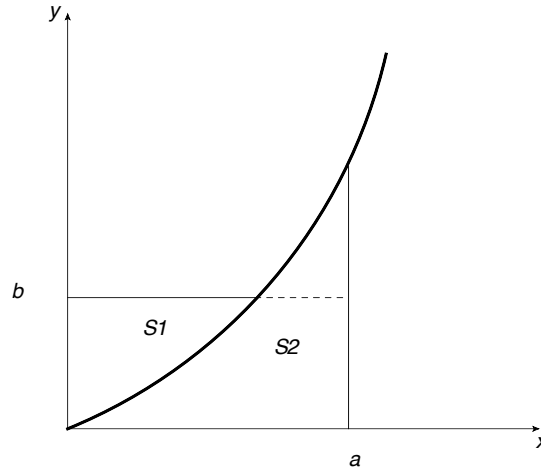


FIGURA 3-4. Dimostrazione grafica della disuguaglianza di Holder.

e ovviamente la somma $S_1 + S_2 \leq ab$. Il segno di uguale vale solo per $b = a^{p-1}$. Ne segue che assumendo senza perdita di generalità[¶] che $\|x\|_p = \|y\|_q = 1$

$$\sum_i |x_i y_i| \leq \sum_i \left(\frac{|x_i|^p}{p} + \frac{|y_i|^q}{q} \right) = \frac{1}{p} + \frac{1}{q} = 1.$$

Ora dall'identità

$$\sum_i (|x_i| + |y_i|)^p = \sum_i ((|x_i| + |y_i|)^{p-1} |x_i| + (|x_i| + |y_i|)^{p-1} |y_i|)$$

segue per la disuguaglianza di Holder

$$\begin{aligned} \sum_i (|x_i| + |y_i|)^p &\leq \|x\|_p \left(\sum_i (|x_i| + |y_i|)^{(p-1)q} \right)^{1/q} + \\ &\quad \|y\|_p \left(\sum_i (|x_i| + |y_i|)^{(p-1)q} \right)^{1/q}. \end{aligned}$$

Ma $q(p-1) = p$ e $1-1/q = 1/p$ per cui

$$\begin{aligned} \sum_i (|x_i| + |y_i|)^p &\leq \left(\sum_i (|x_i| + |y_i|)^p \right)^{1/q} (\|x\|_p + \|y\|_p) \\ &\Downarrow \\ \left(\sum_i (|x_i| + |y_i|)^p \right)^{1/p} &\leq \|x\|_p + \|y\|_p \end{aligned}$$

[¶]Se x e y soddisfano la disuguaglianza di Holder, anche λx e μy la soddisfano per qualunque λ, μ .

Essendo $|x + y| \leq |x| + |y|$, segue anche la disuguaglianza di Minkowski. ■

DISUGUAGLIANZA DI YOUNG. La disuguaglianza $a b \leq a^p/p + b^q/q$ è un caso particolare della disuguaglianza di Young. Sia $f(x)$ una funzione convessa per $x > 0$ ossia

$$(f(x) + f(y))/2 \geq f(\tfrac{1}{2}(x + y)).$$

Non è necessario richiedere che $f(x)$ abbia derivata continua; tuttavia se $f(x)$ è derivabile due volte, la convessità equivale a $f''(x) \geq 0$. Definiamo ora la trasformata di Legendre di $f(x)$:

$$\phi(y) = \max_{x>0} (x y - f(x))$$

Se f è derivabile, allora si ha $y - f'(x) = 0$ e detta $x = \Phi(y)$ la relazione inversa, segue $\phi(y) = y \Phi(y) - f(\Phi(y))$. La trasformazione è utilizzata in meccanica analitica: f è la Lagrangiana, $x = \dot{q}$ e $y = \partial L / \partial \dot{q}$ è il momento coniugato; allora $\phi(y)$ definisce la Hamiltoniana.


Ad esempio, se $f(x) = x^p/p$, $p > 1$, segue $\phi(y) = y^q/q$ con $q = p/(p - 1)$, come è facile verificare. Val la pena di osservare che se la funzione $f(x)$ è asintoticamente $f \approx m x + r$, allora $\phi(y)$ è definita solo per $y < m$, come nel caso della lagrangiana relativistica $f(x) = \sqrt{x^2 + 1} - 1 \Rightarrow \phi(y) = 1 - \sqrt{1 - y^2}$. La disuguaglianza di Young afferma che

$$f(x) + \phi(y) \geq x y$$

Young

il limite inferiore essendo raggiunto per $y = f'(x)$. La dimostrazione è analoga a quella del caso particolare $f = x^p/p$ e verrà omessa. Riferimenti importanti per tutti gli argomenti di questa appendice sono [Arn78, HLP52].

Problemi

 Il simboli \star e $\star\star$ indicano gli esercizi più impegnativi. Il simbolo $\#$ segnala che la soluzione, almeno parziale, è riportata in fondo al capitolo. Importante: Il segno $n \neq \infty$ sarà utilizzato per segnalare argomenti dove la dimensione finita è cruciale e non è possibile estrapolare a spazi ∞ -dimensionali. Una collezione molto ricca di esercizi si trova su [GL74].

PROB 1. Sia (e_1, e_2, \dots, e_n) una base in uno spazio lineare \mathfrak{R} . Dimostrare che l'insieme \mathfrak{R}' dei vettori che soddisfano la relazione $x = \sum_{i=1}^n c_i e_i$ con

$$\begin{cases} a_{11} c_1 + a_{12} c_2 + \dots + a_{1n} c_n = 0 \\ a_{21} c_1 + a_{22} c_2 + \dots + a_{2n} c_n = 0 \\ \vdots \\ a_{m1} c_1 + a_{m2} c_2 + \dots + a_{mn} c_n = 0 \end{cases}$$

$(m \leq n)$ forma un sottospazio lineare di \mathfrak{R} . Qual'è la dimensione di \mathfrak{R}' ?

PROB 2. Siano (e_1, e_2, e_3) vettori ortonormali. Dimostrare che i vettori (e'_1, e'_2, e'_3) definiti da

$$\begin{cases} e'_1 = \frac{1}{\sqrt{2}} e_1 + \frac{i}{\sqrt{2}} e_2 \\ e'_2 = \frac{i}{\sqrt{6}} e_1 + \frac{1}{\sqrt{6}} e_2 + \sqrt{\frac{2}{3}} i e_3 \\ e'_3 = \frac{1}{\sqrt{3}} e_1 + \frac{1}{\sqrt{3}} e_2 - \frac{i}{\sqrt{3}} e_3 \end{cases}$$

sono linearmente indipendenti.

PROB 3. Dimostrare che se i vettori di una base soddisfano la relazione

$$\sum_{j=1}^n |j\rangle \langle j| = \mathbb{1}$$

allora necessariamente $\langle i|j\rangle = \delta_{ij}$.

PROB 4.[#] Calcolare il determinante di Vandermonde

$$V(x) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & \dots & \dots & \dots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}$$

PROB 5.[#] Dimostrare la formula (di Cauchy)

$$\det \left| \frac{1}{1 - x_i y_j} \right|_{i,j=1,\dots,n} = \frac{V(x) V(y)}{\prod_i \prod_j (1 - x_i y_j)} .$$

PROB 6. Sia definita la matrice A :

$$A_{ij} = \begin{cases} x_i (1 - x_j) & (i < j) \\ x_j (1 - x_i) & (i \geq j) \end{cases}$$

dove i numeri reali x_i sono ordinati $0 < x_n < x_{n-1} < \dots < x_2 < x_1 < 1$. Dimostrare che $\det(A) = (1 - x_1) \prod_{j=1}^{n-1} (x_j - x_{j+1}) x_n$.

PROB 7. Determinare la decomposizione spettrale dell'operatore rappresentato dalla matrice

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

PROB 8.# Calcolare il determinante $n \times n$

$$D_n(q) = \det \| q^{(j-k)^2} \|_{j,k=0,1,\dots,n-1}$$

essendo q un numero complesso qualunque.

PROB 9.# Calcolare il determinante

$$\det \begin{vmatrix} 1-z & 2 & 3 & 4 & \dots & n \\ 2 & 3-z & 4 & 5 & \dots & n+1 \\ 3 & 4 & 5-z & 6 & \dots & n+2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ n & n+1 & n+2 & \dots & 2n-1 & -z \end{vmatrix}, \quad (z \in \mathbb{C})$$

PROB 10. Sia \mathcal{P}_n lo spazio lineare costituito dai polinomi a coefficienti complessi in una variabile x di grado n . Sia $e_j = x^j$ ($j = 0, 1, 2, \dots, n$) la base naturale dei monomi. Determinare la matrice rappresentativa dei seguenti operatori lineari:

i) $(\mathcal{D}p)(x) = dp(x)/dx$.

ii) $(\mathcal{Q}p)(x) = x dp(x)/dx$.

iii) $(\mathcal{T}p)(x) = p(x+1)$.

e calcolarne lo spettro.

PROB 11. Determinare l'operatore inverso (se esiste) dei seguenti operatori lineari nello spazio \mathcal{P}_n dell'esercizio precedente: i) $\mathcal{D} + \mathbb{1}$; ii) $\mathcal{Q} + \mathbb{1}$; iii) \mathcal{T} .

PROB 12. Nello spazio \mathcal{P}_n dell'Eserc.10 sia definito un prodotto interno in uno dei modi seguenti

a) $\langle p_1 | p_2 \rangle \equiv \frac{1}{2} \int_{-1}^1 \overline{p_1(x)} p_2(x) dx$

b) $\langle p_1 | p_2 \rangle \equiv \int_0^\infty e^{-x} \overline{p_1(x)} p_2(x) dx$

c) $\langle p_1 | p_2 \rangle \equiv \int_{-\infty}^\infty e^{-x^2} \overline{p_1(x)} p_2(x) dx$

i) determinare la norma del vettore $e_j = x^j$, $j = 0, 1, 2, \dots$ nei tre casi;

ii) studiare i funzionali lineari

a) $\phi_1(p) = \int_{-1}^1 x p(x) dx$

b) $\phi_2(x) = \int_0^\infty e^{-x} \left[p(x) + x \frac{dp(x)}{dx} \right] dx$

$$c) \quad \phi_3(x) = \int_{-\infty}^{\infty} e^{-x^2} x \frac{dp(x)}{dx} dx$$

e determinare i corrispondenti vettori associati ad essi dal Teorema 2.4.2.

PROB 13. Riferendosi al Probl.10, determinare $\exp\{t\mathcal{D}\}$ e $\exp\{t\mathcal{Q}\}$.

PROB 14. Sia $f(z)$ una funzione analitica in un intorno di $z = 0$. Determinare la matrice rappresentativa dell'operatore $f(\mathcal{D})$ (Probl.10).

PROB 15. ^{*} Sia definito il prodotto interno in \mathcal{P}_n in modo tale che $\langle x^j | x^k \rangle = \binom{n}{j}^{-1} \delta_{jk}$. Si consideri l'operatore

$$(\mathcal{J}p)(x) = \frac{1}{2}nx p(x) + \frac{1}{2}(1-x^2) dp(x)/dx.$$

Determinare la decomposizione spettrale di \mathcal{J} . Verificare che gli autovettori di \mathcal{J} sono ortogonali tra loro.

PROB 16. Determinare lo spettro delle seguenti matrici

$$A = \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix} \quad B = \begin{pmatrix} 3x & \sqrt{3} & 0 & 0 \\ \sqrt{3} & x & 2 & 0 \\ 0 & 2 & -x & \sqrt{3} \\ 0 & 0 & \sqrt{3} & -3x \end{pmatrix}$$

Spiegare per quale motivo gli autovalori sono proporzionali: $E_B = \sqrt{1+x^2} E_A$.

PROB 17. Calcolare il logaritmo dell'operatore rappresentato dalla matrice

$$\begin{pmatrix} \cos \alpha & i \sin \alpha \\ i \sin \alpha & \cos \alpha \end{pmatrix}.$$

PROB 18. Discutere l'esistenza e la molteplicità delle soluzioni dell'equazione $\mathcal{X}^2 = \mathcal{A}$, essendo \mathcal{A} un assegnato operatore lineare.

PROB 19. Sia $\mathcal{X}_{jk} = e^{aj+bk}$, con $a, b \in \mathbb{C}$, $j, k = 1, 2, \dots, n$. Determinare la decomposizione spettrale della matrice \mathcal{X}

PROB 20.[#] Siano \mathcal{A} e \mathcal{B} due operatori lineari. Dimostrare che gli operatori $\mathcal{A}\mathcal{B}$ e $\mathcal{B}\mathcal{A}$ hanno lo stesso polinomio caratteristico (e lo stesso polinomio minimale?).

PROB 21.[#] Dimostrare le formule di Faddeev per il polinomio caratteristico

$$\Delta(\lambda) = \det(\lambda \mathbb{1} - A) \equiv \lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \dots - p_{n-1} \lambda - p_n$$

$$\begin{array}{lll} \mathcal{A}_1 \equiv A B_0 & p_1 = \text{Tr}(A_1) & B_1 \equiv A_1 - p_1 \mathbb{1}, \quad (B_0 \equiv \mathbb{1}) \\ \dots & \dots & \dots \\ \mathcal{A}_k \equiv A B_{k-1} & p_k = \frac{1}{k} \text{Tr}(A_k) & B_k \equiv A_k - p_k \mathbb{1}, \quad (1 \leq k \leq n) \\ \dots & \dots & \dots \\ \mathcal{A}_n \equiv A B_{n-1} & p_n = \frac{1}{n} \text{Tr}(A_n) & B_n \equiv A_n - p_n \mathbb{1} \equiv 0 \end{array}$$

PROB 22. *Dimostrare che se*

$$\det(z\mathbb{1} - A) = \prod_1^m (z - \lambda_k)^{r_k}$$

allora

$$\det(z\mathbb{1} - A^2) = \prod_1^m (z - \lambda_k^2)^{r_k}$$

PROB 23. * *Dimostrare che per qualunque $f(z)$ tale che esista $\hat{f}(A)$ si ha (con le notazioni del Problema precedente)*

$$\det(z\mathbb{1} - \hat{f}(A)) = \prod_1^m (z - f(\lambda_k))^{r_k}$$

PROB 24. # *Dimostrare l'identità*

$$\det(e^A) = e^{\text{Tr}(A)}$$

essendo $\text{Tr}(A)$ la traccia di A , ossia la somma degli elementi di matrice sulla diagonale.

PROB 25. *Dimostrare che i proiettori spettrali e i nilpotenti che entrano nella decomposizione spettrale di un operatore lineare soddisfano le relazioni*

- i) $\text{Tr}(E_\lambda^m) = 0$ per $m > 0$
- ii) $\text{Tr}(P_\lambda) = r_\lambda$ ($r_\lambda = \text{molteplicità di } \lambda$).

PROB 26. #** *La matrice U di dimensione $n+1$ è definita da*

$$U_{jk} = (n+1)^{-1/2} \exp\{2\pi i j k / (n+1)\}$$

($j, k = 0, 1, 2, \dots, n$). Dimostrare che U è unitaria e calcolarne lo spettro. (U rappresenta la tabella dei caratteri irriducibili del gruppo delle permutazioni S_{n+1} , alias la matrice che realizza la trasformata di Fourier discreta).

PROB 27. *Calcolare lo spettro dell'operatore $B(t)$ al secondo ordine in teoria delle perturbazioni:*

$$B(t) \leftarrow \begin{pmatrix} 2 & t & 0 & 0 & 0 \\ t & 1 & t\sqrt{\frac{3}{2}} & 0 & 0 \\ 0 & t\sqrt{\frac{3}{2}} & 0 & t\sqrt{\frac{3}{2}} & 0 \\ 0 & 0 & t\sqrt{\frac{3}{2}} & -1 & t \\ 0 & 0 & 0 & t & -2 \end{pmatrix}$$

PROB 28. *Sia $A \leftarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. Determinare l'operatore $\exp\{itA\}$.*

PROB 29. *L'operatore \mathcal{F} è rappresentato dalla matrice*

$$\mathcal{F} \leftarrow \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

- i) *determinare la decomposizione spettrale di \mathcal{F} ;*
- ii) *calcolare $V(t) = \exp(t\mathcal{F})$*

iii) calcolare $\det(V(t))$.

PROB 30. Sia X l'operatore definito nel Probl.19. Quali condizioni devono essere imposte su a, b affinché \mathcal{X} sia i) normale, ii) autoaggiunto ?

PROB 31.*** Siano a_1, a_2, \dots, a_n gli autovalori della matrice A di dimensione n , e siano T_k le tracce $T_k = \text{Tr}(A^k) = \sum_j a_j^k$. Dimostrare che il polinomio caratteristico $\Delta(z)$ della matrice A ha la forma seguente

$$\Delta(z) \equiv \det(z\mathbb{1} - A) = \sum_0^n p_j(A) z^{n-j}$$

dove i coefficienti sono determinati dalle formule di Newton

$$p_j(A) = \frac{1}{j!} \det \begin{vmatrix} T_1 & 1 & 0 & 0 & \dots & 0 \\ T_2 & T_1 & 2 & 0 & \dots & 0 \\ T_3 & T_2 & T_1 & 3 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \dots & \vdots \\ T_{j-1} & T_{j-2} & \dots & T_2 & T_1 & j-1 \\ T_j & T_{j-1} & \dots & T_3 & T_2 & T_1 \end{vmatrix}$$

PROB 32.*** Con le stesse notazioni del problema precedente, dimostrare che vale lo sviluppo in serie (nota come formula di Brioschi)

$$\det(1 - tA)^{-1} = \sum_{j=0}^{\infty} q_j(A) t^j$$

dove

$$q_j(A) = \frac{1}{j!} \det \begin{vmatrix} T_1 & -1 & 0 & 0 & \dots & 0 \\ T_2 & T_1 & -2 & 0 & \dots & 0 \\ T_3 & T_2 & T_1 & -3 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \dots & \vdots \\ T_{j-1} & T_{j-2} & \dots & T_2 & T_1 & -j+1 \\ T_j & T_{j-1} & \dots & T_3 & T_2 & T_1 \end{vmatrix}$$

PROB 33. ** Secondo le notazioni del Probl. 31, ricavare le seguenti identità valide per il caso $a_j = a q^{j-1}$, $q \neq 1$:

$$\Delta(z) = \sum_{j=0}^n \begin{bmatrix} n \\ j \end{bmatrix}_q (-a)^j q^{j(j-1)/2} z^{n-j}$$

$$\Delta(z)^{-1} = \sum_{j=0}^{\infty} \begin{bmatrix} n+j-1 \\ j \end{bmatrix}_q a^j z^{-n-j}$$

dove

$$\begin{bmatrix} n \\ j \end{bmatrix}_q = \frac{(q; q)_n}{(q; q)_j (q; q)_{n-j}}, \quad (\alpha, q)_n \equiv \prod_{j=1}^n (1 - \alpha q^{j-1}), \quad [n]! = (q, q)_n$$

PROB 34. # Sia \mathcal{A}_0 un operatore autoaggiunto invertibile e sia $\mathbb{P} = |\mathbf{v}\rangle \langle \mathbf{v}|$. Dimostrare che l'operatore $\mathcal{A} = \mathcal{A}_0 + \varepsilon \mathbb{P}$ è invertibile per ε sufficientemente piccolo e il suo inverso è dato da

$$\mathcal{A}^{-1} = \mathcal{A}_0^{-1} + \varepsilon \frac{\mathcal{A}_0^{-1} \mathbb{P} \mathcal{A}_0^{-1}}{\mathbb{1} + \varepsilon \text{Tr}(\mathcal{A}_0^{-1} \mathbb{P})}.$$

PROB 35. Con riferimento all'esercizio precedente, dimostrare che

$$\det(z\mathbb{1} - A_0 - \varepsilon \mathbb{P}) = \det(z\mathbb{1} - A_0) \left(1 - \varepsilon \langle v | (z\mathbb{1} - A_0)^{-1} | v \rangle\right).$$

Questa relazione si può utilizzare per calcolare lo spettro di A noto quello di A_0 .

PROB 36. Sia Q la matrice $Q_{ij} = i + j - 1$, ($i, j = 1, 2, \dots, n$). Dimostrare che vale l'equazione

$$Q^3 - n^2 Q^2 - \frac{1}{12} n^2 (n^2 - 1) Q = 0$$

e dedurne la decomposizione spettrale (vedi anche il Probl.9).

PROB 37. Sia \mathcal{P}_n lo spazio dei polinomi complessi con prodotto interno come nel Probl. 15. Determinare la rappresentazione matriciale dell'operatore $(\mathcal{D}p)(x) = dp(x)/dx$ nella base ortonormale $e_j = \binom{n}{j}^{1/2} x^j$, ($j = 0, 1, 2, \dots, n$). Costruire l'aggiunto Hermitiano di \mathcal{D} e darne una rappresentazione indipendente dalla base.

PROB 38.[#] (Modello di Eherenfest). Sono date due urne contenenti N biglie numerate da 1 a N . Si estrae ripetutamente un numero a caso compreso tra 1 e N e la biglia corrispondente al numero estratto cambia posto dall'urna in cui si trova all'altra*.

a) Dimostrare che la probabilità $p(n, t)$ di trovare n biglie nella prima urna dopo t estrazioni soddisfa la relazione

$$\begin{pmatrix} p(0, t+1) \\ p(1, t+1) \\ p(2, t+1) \\ \vdots \\ p(N-2, t+1) \\ p(N-1, t+1) \\ p(N, t+1) \end{pmatrix} = \begin{pmatrix} 0 & 1/N & 0 & 0 & \dots & 0 \\ 1 & 0 & 2/N & 0 & 0 & \dots & 0 \\ 0 & 1-1/N & 0 & 3/N & 0 & \dots & 0 \\ \vdots & \dots & \ddots & \dots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 3/N & 0 & 1-1/N \\ 0 & 0 & \dots & 2/N & 0 & 1 \\ 0 & 0 & 0 & \dots & 1/N & 0 \end{pmatrix} \begin{pmatrix} p(0, t) \\ p(1, t) \\ p(2, t) \\ \vdots \\ p(N-2, t) \\ p(N-1, t) \\ p(N, t) \end{pmatrix}$$

ossia

$$p(n, t+1) = \frac{1}{N} (N - n + 1) p(n - 1, t) + \frac{1}{N} (n + 1) p(n + 1, t).$$

- b) Sfruttando il risultato del Probl. 15 determinare la probabilità di avere un'urna vuota dopo t secondi se inizialmente la stessa urna era vuota;
c) Scrivere un programma **matlab** per simulare il processo.

PROB 39. Sia $\mathcal{A}(t)$ un operatore lineare dipendente da un parametro reale t . Supponendo che \mathcal{A} sia derivabile, dimostrare che la derivata dell'operatore inverso è data da

$$\frac{d}{dt} \mathcal{A}^{-1}(t) = -\mathcal{A}^{-1}(t) \frac{d\mathcal{A}(t)}{dt} \mathcal{A}^{-1}(t)$$

PROB 40. Determinare la soluzione del sistema lineare

$$\begin{cases} \frac{dx}{dt} = -x + y \\ \frac{dy}{dt} = -y + z \\ \frac{dz}{dt} = w - 3x \\ \frac{dw}{dt} = -z + y \end{cases}$$

essendo $x = 1, y = 1, z = 1, w = 3$ per $t = 0$.

*Per apprezzare appieno questo esercizio è meglio riprenderlo quando si sia studiata la Parte IV.

PROB 41. Calcolare l'esponenziale $\exp(tZ)$ essendo Z rappresentato dalla matrice

$$Z \leftrightarrow \begin{pmatrix} 0 & n_3 & -n_2 \\ -n_3 & 0 & n_1 \\ n_2 & -n_1 & 0 \end{pmatrix}$$

con $n_1^2 + n_2^2 + n_3^2 = 1$.

PROB 42. Si determini lo spettro della matrice $A(t)$ al secondo ordine in teoria delle perturbazioni e confrontare il risultato con quello esatto:

$$A = \begin{pmatrix} 0 & 1 & t & 0 \\ t & 0 & 0 & t \\ t & 0 & 0 & -2i \\ 0 & t & 2i & 0 \end{pmatrix}$$

. Per quali valori di t lo sviluppo perturbativo converge?

PROB 43. Un operatore lineare T ha la seguente rappresentazione matriciale detta "triangolare superiore":

$$T \leftrightarrow \begin{pmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1n} \\ 0 & t_{22} & t_{23} & \dots & t_{2n} \\ 0 & 0 & t_{33} & \dots & t_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & t_{nn} \end{pmatrix}$$

Qual'è lo spettro di T ? Nell'ipotesi che tutti gli autovalori siano distinti calcolare $(\exp T)_{ii}$ e $(\exp T)_{ii+1}$.

PROB 44. Dimostrare che se X ha tutti autovalori distinti e Y commuta con X allora Y è diagonale nella stessa base in cui lo è X .

PROB 45.[#] Determinare autovalori ed autovettori della matrice di dimensione n

$$X = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 1 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

PROB 46.[#] Sia Y una matrice che soddisfa la relazione

$$Y_{j+1k} + Y_{j-1k} = Y_{jk+1} + Y_{jk-1}$$

(matrice iperbolica) con la convenzione che $Y_{ij} = 0$ se un indice è esterno a $[1, n]$.

- i) Dimostrare che Y è individuata univocamente dalla sua prima riga;
- ii) assegnata la prima riga, calcolare lo spettro di Y ;
- iii) diagonalizzare la matrice che si trova in copertina.

PROB 47.[#] La matrice (iperbolica) K di dimensione N è definita da

$$\begin{cases} K_{ii} = (N+1)/2 \\ K_{ij} = 0 \\ K_{ij} = \frac{1}{2(N+1)} \left(\left(\cot \frac{\pi(i+j)}{2(N+1)} \right)^2 - \left(\cot \frac{\pi(i-j)}{2(N+1)} \right)^2 \right) \end{cases} \quad \begin{matrix} (i+j \text{ pari}) \\ \\ (i+j \text{ dispari}) \end{matrix}$$

Calcolarne lo spettro.

PROB 48. #* Sia $U^{(n)}$ una matrice che soddisfa la condizione del Probl. 46, con $U_{ij}^{(n)} = \delta_{j,n+1}$ e $U^{(1)} = X$ la matrice del problema 45. Dimostrare la relazione di ricorrenza

$$\mathcal{X} U^{(n)} = U^{(n-1)} + U^{(n+1)}.$$

Ricavare la forma esplicita della dipendenza funzionale $U^{(n)} = f_n(\mathcal{X})$.

PROB 49. * Determinare esplicitamente la matrice

$$K = \frac{n+1}{\pi} \cos^{-1}(\mathcal{X}/2)$$

dove al solito \mathcal{X} si riferisce alla matrice del problema 45.

PROB 50. #* Dimostrare che la matrice $n \times n$

$$M_{jk} = \begin{cases} \sqrt{(1-x_j^2)(1-x_k^2)} (x_j - x_k)^{-2} & (j \neq k) \\ \sum_{i \neq j} (1-x_i^2) (x_j - x_i)^{-2} & (j = k) \end{cases}$$

dove $x_j = \cos \frac{j\pi}{n+1}$, è funzione di \mathcal{X} (45) e determinarne lo spettro.

PROB 51. Una matrice C è detta "circolante" se le sue righe sono ottenute dalla prima permutando successivamente in modo ciclico gli elementi di matrice, ad es.

$$C = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$$

Sia C una matrice circolante di dimensione N . Dimostrare che lo spettro di C è dato da

$$\lambda_n = \sum_{j=1}^N C_{1j} \omega_n^{j-1}$$

avendo indicato con ω_n l' n -esima radice dell'unità.

PROB 52. # Sia \mathcal{J} l'operatore lineare (nilpotente) rappresentato dalla matrice $\mathcal{J}_{ij} = \delta_{ij+1}$. Dimostrare che ogni operatore che commuta con \mathcal{J} è funzione di \mathcal{J} .

PROB 53 (Coppie di Lax). Siano $A(t)$ e $B(t)$ operatori lineari che soddisfano l'equazione

$$\frac{dA(t)}{dt} = B(t)A(t) - A(t)B(t);$$

dimostrare che lo spettro di A è indipendente dal tempo.

PROB 54. Siano A e B come nel problema precedente. In particolare sia

$$A = \begin{pmatrix} a_1 & b_1 & 0 & b_4 \\ b_1 & a_2 & b_2 & 0 \\ 0 & b_2 & a_3 & b_3 \\ b_4 & 0 & b_3 & a_4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -b_1 & 0 & b_4 \\ b_1 & 0 & -b_2 & 0 \\ 0 & b_2 & 0 & -b_3 \\ -b_4 & 0 & b_3 & 0 \end{pmatrix}$$

Dimostrare che attraverso la trasformazione

$$a_j = \frac{1}{2}p_j, \quad b_j = \frac{1}{2} \exp\{(q_{j-1} - q_j)/2\}, \quad (j = 1, 2, 3, 4)$$

con la convenzione $q_0 \equiv q_4$, l'equazione $dA/dt = [B(t), A(t)]$ si riconduce alle equazioni del moto Hamiltoniane di un sistema classico con coordinate canoniche (q_i, p_i) . Determinare l'Hamiltoniana e generalizzare il risultato a matrici di dimensione qualunque. Nota: dal risultato del

problema precedente, lo spettro di $A(t)$ è indipendente dal tempo; questo comporta che il sistema è integrabile. Determinare le costanti del moto[†].

PROB 55. Dimostrare che per ogni coppia di operatori lineari \mathcal{X}, \mathcal{Y} vale l'identità

$$e^{t\mathcal{X}} e^{s\mathcal{Y}} e^{-t\mathcal{X}} e^{-s\mathcal{Y}} = 1 + ts(\mathcal{X}\mathcal{Y} - \mathcal{Y}\mathcal{X}) + O(t^2).$$

PROB 56 (Matrici tridiagonali).^{#*} Una matrice è detta tridiagonale se ha elementi nulli tranne che per la diagonale principale e le due diagonali adiacenti, cioè $T_{ij} = 0$ per $|i - j| > 1$. Sia T una matrice tridiagonale reale e simmetrica. Dimostrare che se tutti gli elementi $T_{i,i+1}$ sono diversi da zero, allora tutti gli autovalori di T sono semplici.

PROB 57. Calcolare lo spettro della matrice circolante

$$\mathcal{D} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \cdot & \ddots & & \ddots & & \vdots \\ \vdots & & & \ddots & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & 1 & 0 & 1 \\ 1 & 0 & 0 & \dots & \dots & 1 & 0 \end{pmatrix}$$

PROB 58.^{#**} Dimostrare, facendo uso della soluzione del Cap.3.5.3, la formula di Christoffel–Darboux

$$\sum_{n=0}^{N-1} p_n(x) p_n(y) = \frac{\kappa_{N-1}}{\kappa_N} \frac{p_N(x) p_{N-1}(y) - p_{N-1}(x) p_N(y)}{x - y}$$

dove $p_n(x) = \kappa_n x^n + \dots$ è una base di polinomi ortonormali.

PROB 59. Sia \mathcal{P}_N lo spazio dei polinomi ortogonali di grado $\leq N$ rispetto a una qualunque misura di integrazione. Dimostrare che la funzione

$$K(x, y) \equiv \sum_{n=0}^N \overline{\phi_n(x)} \phi_n(y)$$

definita per una scelta di vettori di base ortonormali $\phi_n(x)$ è indipendente dalla base. Inoltre mostrare che per ogni polinomio $p(x)$ vale la relazione

$$\langle e_x | p \rangle = p(x)$$

essendo e_x il polinomio $e_x(y) = K(x, y)$. In altre parole, $K(x, y)$ fa le funzioni della $\delta(x, y)$ di Dirac nello spazio \mathcal{P}_N .

PROB 60.[#] Fissato un punto x_0 si determini il polinomio $p(x)$ di norma 1 in \mathcal{P}_N tale che risulti massimo il valore $p(x_0)$.

PROB 61. Sia definito il prodotto interno nello spazio \mathcal{P}_N nel modo seguente

$$\langle p_1 | p_2 \rangle = \int_{\mathbb{R}^2} dx dy \frac{\overline{p_1(x + iy)} p_2(x + iy)}{(1 + x^2 + y^2)^{N+2}}$$

[†]Questo problema è un caso particolare del cosiddetto “modello di Toda”, un modello di sistemi classici integrabili a n -gradi di libertà. Vedere i contributi di Flaschka e Moser in [EHW75].

- i) Determinare la normalizzazione dei vettori $e_k(x) = \mathcal{N}_k z^k$ in modo che formino una base ortonormale ($z = x + iy$).
- ii) Calcolare la funzione $K(x, y)$ definita nel Probl. 59.
- iii) Dimostrare che per $\alpha, \beta \in \mathbb{C}$, $|\alpha|^2 + |\beta|^2 = 1$, l'operatore $\mathcal{U}(\alpha, \beta)$ definito come

$$(\mathcal{U}p)(z) = (-\bar{\beta}z + \bar{\alpha})^N p\left(\frac{\alpha z + \beta}{-\bar{\beta}z + \bar{\alpha}}\right)$$

è unitario.

- iv) Selezionati $\alpha = \cos \frac{1}{2}\Phi$, $\beta = \sin \frac{1}{2}\Phi$, si determini l'operatore \mathcal{J} tale che $\mathcal{U} = \exp(i\Phi\mathcal{J})$.

PROB 62. Dimostrare l'identità

$$\text{ad}_X^n(Y) = \sum_{j=0}^n \binom{n}{j} X^j Y (-X)^{n-j}$$

essendo $\text{ad}_X(Y) \equiv [X, Y]$. Suggerimento: applicare il principio di induzione matematica.

PROB 63.[#] Un dado da gioco, di cui si mostra lo sviluppo delle facce nella Fig. 3-1, presenta la numerazione da 1 a 6. Se si sostituisce ad ogni numero la media dei primi vicini (es. $5 \rightarrow (1 + 2 + 3 + 4)/4$) si ottiene un dado con numerazione razionale. Se il procedimento è ripetuto indefinitamente quale sarà il valore limite su ogni faccia? (Per una soluzione che richiede nozioni di teoria dei gruppi si veda [Kir74] §16). Generalizzare il problema a un poliedro regolare.

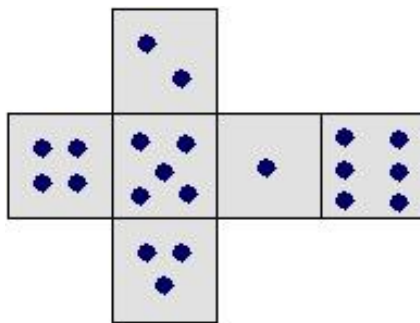


FIGURA 3-1. Il dado del Probl. 63

PROB 64. Un operatore lineare \mathcal{A} soddisfa l'equazione minimale $\mathcal{A}^3(1 - \mathcal{A})^2 = 0$. Determinare la decomposizione spettrale di \mathcal{A} .

PROB 65. Risolvere il sistema di equazioni differenziali

$$\frac{d}{dt} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

con le condizioni iniziali $(x, y, z, w)_{t=0} = (1, 0, 0, 0)$.

PROB 66.[#] Sia \mathcal{K} un operatore tale che $\mathbb{1} - \mathcal{K}$ sia invertibile, esiste cioè unica la soluzione dell'equazione $x = \mathcal{K}x + y$, con y assegnato.

i) Discutere la convergenza del procedimento iterativo

$$\mathbf{x}_{t+1} = \mathcal{K} \mathbf{x}_t + \mathbf{y}, \quad \mathbf{x}_0 = 0, \quad t = 1, 2, 3, \dots$$

ii) mostrare che, scelto un parametro reale ω compreso tra 0 e 1, il nuovo procedimento iterativo

$$\mathbf{x}_{t+1} = (1 - \omega) \mathbf{x}_t + \omega (\mathcal{K} \mathbf{x}_t + \mathbf{y}), \quad \mathbf{x}_0 = 0, \quad t = 1, 2, 3, \dots$$

se converge determina la stessa soluzione.

Il parametro ω è noto come parametro di rilassamento e in genere facilita la convergenza dell'iterazione.

PROB 67. Sia $\mathcal{R}(z) = (z\mathbb{1} - \mathcal{A})^{-1}$ il risolvente di \mathcal{A} . Dimostrare che se z_0 è un autovalore di \mathcal{A} allora $\mathcal{R}(z)$ ammette lo sviluppo (di Laurent)

$$\mathcal{R}(z) = \sum_{n=-\nu-1}^{\infty} A_n (z - z_0)^n,$$

e dimostrare che $A_{-2} A_{-1} = A_{-2}$, $A_{-1}^2 = A_{-1}$.

PROB 68. Sia $\mathcal{P}_n^{(2)}$ lo spazio dei polinomi in due variabili di grado $\leq n$.

iii) Si determini lo spettro degli operatori

$$\begin{aligned} \mathcal{M} &= i \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \\ \mathcal{R} &= x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}. \end{aligned}$$

iv) Dimostrare che il commutatore $[\mathcal{M}, \mathcal{R}]$ è zero e sfruttare questo fatto per semplificare il calcolo degli autovalori di \mathcal{M} .

v) Siano $\mathcal{K}_1 = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}$, $\mathcal{K}_2 = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}$; determinare i commutatori tra $\mathcal{K}_1, \mathcal{K}_2, \mathcal{M}$.

PROB 69.# Sia $\mathcal{H}_\ell^{(3)}$ lo spazio dei polinomi armonici in tre variabili di grado $\ell \in \mathbb{Z}^+$:

$$\begin{aligned} \mathbf{p} \in \mathcal{H}_\ell^{(3)} : \quad \mathbf{p}(x, y, z) &= \sum_{i+j+k=\ell} c_{ijk} x^i y^j z^k \\ \Delta \mathbf{p} &\equiv \left[\left(\frac{\partial}{\partial x} \right)^2 + \left(\frac{\partial}{\partial y} \right)^2 + \left(\frac{\partial}{\partial z} \right)^2 \right] \mathbf{p}(x, y, z) = 0. \end{aligned}$$

Determinare lo spettro dell'operatore $\mathcal{M}_3 = i \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)$.

PROB 70. Sia

$$\mathbf{u} \in \text{SU}(2) \leftrightarrow \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

una matrice unitaria 2×2 . Sia inoltre $\mathbf{X} = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}$ una matrice Hermitiana a traccia nulla, pure 2×2 . Si dimostri che $\mathbf{X}' = \mathbf{u} \mathbf{X} \mathbf{u}^\dagger$ è Hermitiana a traccia nulla e che gli elementi di matrice di \mathbf{X}' definiscono le variabili x', y', z' legate a x, y, z da una rotazione spaziale.



PROB 71. Sia S una generica matrice complessa 2×2 con determinante uguale a uno ($S \in \text{SL}(2, \mathbb{C})$). Sia inoltre $X = \begin{pmatrix} z+t & x-iy \\ x+iy & -z+t \end{pmatrix}$ con x, y, z, t reali. Si verifichi che $X' = SXS^\dagger$ è della stessa forma di X con variabili x', y', z', t' legate a x, y, z, t da una trasformazione di Lorentz.

PROB 72. $n \neq \infty$ Dimostrare che la serie definita dalla Eq.3.13 converge per ogni t e ogni coppia di operatori \mathcal{A}, \mathcal{B} .

PROB 73. $n \neq \infty$ Dimostrare che le relazioni


$$[\mathcal{A}, \mathcal{B}] = \mathcal{Z}, [\mathcal{A}, \mathcal{Z}] = \mathbb{O}, [\mathcal{B}, \mathcal{Z}] = \mathbb{O},$$

implicano $\mathcal{Z} = \mathbb{O}$ nel caso che \mathcal{A} sia diagonalizzabile.

 In Meccanica Quantistica si considerano gli operatori canonici q e p che soddisfano le regole di commutazione di Heisenberg $[q, p] = i\hbar \mathbb{1}$. La tesi del problema implica che q e p non possono essere realizzati con matrici hermitiane finito-dimensionali. In realtà vale un teorema più generale. Si tornerà su questo nella Parte III (§ 4.3.3). 

PROB 74. $n \neq \infty$ Siano A e B due operatori lineari tali che $A+tB$ abbia spettro indipendente da t . Dimostrare che B è nilpotente.

PROB 75. $n \neq \infty$ Siano A, B e C operatori lineari tali che $A+tB+t^2C$ abbia spettro indipendente da t . Dimostrare che C è nilpotente.

 Anche questo risultato non si estende a infinite dimensioni, ad es. per gli operatori q, p della Meccanica Quantistica vale la relazione

$$e^{-itp/\hbar} (p^2 + q^2) e^{itp/\hbar} = (p^2 + q^2) - 2tq + t^2 \mathbb{1}$$

che viola palesemente l'affermazione.

PROB 76. $n \neq \infty$ L'operatore $\mathbb{1}$ non può essere un commutatore, cioè non esistono operatori \mathcal{A}, \mathcal{B} tali che $[\mathcal{A}, \mathcal{B}] = \mathbb{1}$.

PROB 77. Determinare lo spettro dell'operatore

$$X(a) = \begin{pmatrix} 0 & 1 & a & 0 & \dots & 0 \\ 1 & a & 1 & a & 0 & \dots & 0 \\ a & 1 & a & 1 & a & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & a & 1 & a & 1 & a \\ 0 & \dots & 0 & a & 1 & a & 1 \\ 0 & \dots & 0 & a & 1 & 0 \end{pmatrix}$$

corrispondente a una matrice definita secondo il Probl. 46 utilizzando la teoria delle perturbazioni nel parametro a .

PROB 78. n Siano $|a_1\rangle, \dots, |a_r\rangle$ e $|b_1\rangle, \dots, |b_r\rangle$ vettori arbitrari in uno spazio n -dimensionale ($r < n$). Dimostrare che lo spettro dell'operatore

$$\mathcal{A} = \sum_j |a_j\rangle \langle b_j|$$

è dato dalle r radici dell'equazione

$$\det |z\delta_{ij} - \langle b_i | a_j \rangle| = 0$$

mentre i rimanenti $n - r$ autovalori sono tutti nulli.

PROB 79. Una matrice di dimensione n ha tutti elementi nulli al di fuori della prima riga e della prima colonna. Trovare il suo spettro.

PROB 80.[#] Una matrice di dimensione n ha tutti elementi nulli al di fuori delle prime due righe e delle prime due colonne. Ridurre il problema della determinazione dello spettro a quello di una matrice 4×4 .

PROB 81.^{#*} La matrice rappresentativa di $\mathcal{H}(\lambda)$ è data da

$$\mathcal{H} \leftarrow \begin{pmatrix} 1 & \lambda & 0 & \dots & \dots & \dots & 0 \\ \lambda & 2 & \lambda & 0 & \dots & \dots & 0 \\ 0 & \lambda & 3 & \lambda & 0 & \dots & 0 \\ 0 & 0 & \lambda & 4 & \lambda & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda & 5 & \lambda & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \lambda & 6 & \lambda & 0 & \dots & 0 \\ \vdots & & \dots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \dots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & \lambda & N-1 & \lambda \\ 0 & \dots & \dots & \dots & 0 & \lambda & N \end{pmatrix}$$

Dimostrare che l' n -esimo autovalore ($n < N/2$) ammette uno sviluppo in serie del tipo $E_n = n + \delta(n)\lambda^{2n}$ e calcolare $\delta(n)$.

PROB 82. ^{*} Vale la seguente rappresentazione per lo sviluppo di Baker–Campbell–Hausdorff (Rychtmeyer e Greenspan, *Commun. Pure Appl. Math.* XVIII (1965) pag. 107):

$$Z = \log(e^X e^Y) = X + \int_0^1 \Psi(e^{\text{ad}_X} e^{t \text{ad}_Y}) Y dt$$

dove

$$\Psi(z) = \frac{z \log z}{z - 1}$$

Utilizzare questa formula per ottenere i primi termini dello sviluppo come nell'Eq.3.14.

Soluzioni

#4 Il determinante è un polinomio omogeneo di grado $\frac{1}{2}n(n-1)$ nelle variabili x_i e risulta divisibile per $\prod_{i<j}(x_i - x_j)$ in quanto si annulla se per una scelta qualunque di i, j si ha $x_i = x_j$. Ma il prodotto è un polinomio omogeneo dello stesso grado del determinante e quindi il quoziente è una costante che si trova facilmente essere uguale a uno.

#5 Moltiplichiamo il determinante per $\prod_{i,j}(1 - x_i y_j)$ e distribuiamo i fattori contenenti x_i alla i -esima riga. Ne risulta il determinante

$$\det |\prod_{i \neq j}(1 - x_i y_j)|$$

che è un polinomio di grado $n-1$ in tutte le variabili x_i e y_i . Dato che esso si annulla se per una coppia qualunque $x_i = x_j$, esso deve contenere il fattore $V(x)$ (Vandermonde) e analogamente $V(y)$. Contando le potenze si vede che il fattore che rimane indeterminato deve essere una costante, che si determina controllando il limite $x_n \rightarrow \infty$.

#8 Il determinante è un Vandermonde camuffato!

#9 La matrice che si ottiene ponendo $z = 0$ è di rango 2, infatti la somma della riga $i+1$ con la riga $i-1$ è il doppio della riga i -esima. Questo implica che nello sviluppo del determinante caratteristico compaiono soltanto le potenze z^n, z^{n-1} e z^{n-2} . Si trova quindi

$$\det = (-z)^n + (-z)^{n-1}(1 + 3 + 5 + \dots + 2n - 1) + (-z)^{n-2} \sum_{i < j} ((2i-1)(2j-1) - (i+j-1)^2) = \text{etc.}$$

#20 L'affermazione è ovvia se uno dei due operatori è invertibile. Infatti in questo caso si ha $\det(\lambda \mathbb{1} - \mathcal{A}\mathcal{B}) = \det(\mathcal{A}(\lambda \mathbb{1} - \mathcal{B}\mathcal{A})\mathcal{A}^{-1}) = \det(\lambda \mathbb{1} - \mathcal{B}\mathcal{A})$. Inoltre la decomposizione spettrale di $\mathcal{A}\mathcal{B}$ si ottiene per trasformazione di similitudine da quella di $\mathcal{B}\mathcal{A}$, e quindi gli indici ν_λ sono indipendenti dall'ordine degli operatori. Se entrambi gli operatori sono singolari, allora consideriamo il determinante in cui abbiamo perturbato $\mathcal{A} \rightarrow \mathcal{A} + \varepsilon \mathbb{1}$; per un intorno di $\varepsilon = 0$ ci si riduce al caso precedente: segue

$$\det(\lambda \mathbb{1} - (\mathcal{A} + \varepsilon \mathbb{1})\mathcal{B}) = \det(\lambda \mathbb{1} - \mathcal{B}(\mathcal{A} + \varepsilon \mathbb{1}))$$

e i due determinanti sono polinomi in ε in cui possiamo passare al limite $\varepsilon \rightarrow 0$ con continuità. Il risultato è anche conseguenza immediata dello sviluppo di Newton del polinomio caratteristico (vedi Probl. 31) in quanto esprimibile in termini di tracce; si ha $\text{Tr}((\mathcal{A}\mathcal{B})^k) = \text{Tr}((\mathcal{B}\mathcal{A})^k)$. Il polinomio minimale invece **non** è necessariamente lo stesso. I sottospazi corrispondenti ad autovalori $\lambda \neq 0$ non costituiscono un problema. Infatti da

$$\mathcal{A}\mathcal{B}\psi_j = \lambda \psi_j, \quad (j = 1, 2, \dots, r_\lambda - \nu_\lambda)$$

segue

$$\mathcal{B}\mathcal{A}\mathcal{B}\psi_j = \lambda \mathcal{B}\psi_j, \quad (j = 1, 2, \dots, r_\lambda - \nu_\lambda)$$

e i vettori $\mathcal{B}\psi_j$ sono linearmente indipendenti. Dunque gli indici di $\mathcal{A}\mathcal{B}$ e $\mathcal{B}\mathcal{A}$ corrispondenti a $\lambda \neq 0$ coincidono e di conseguenza i polinomi minimali coincidono limitatamente allo spettro di autovalori non nulli. Per l'autovalore $\lambda = 0$ si ha

$$\dim \ker(\mathcal{A}\mathcal{B}) = \dim \ker(\mathcal{B}\mathcal{A})$$

($\ker(\mathcal{X})$ denota il sottospazio $\mathcal{X}\psi = 0$) in quanto coincidono le dimensioni dei sottospazi relativi a $\lambda \neq 0$, ma gli indici $\nu_0(\mathcal{AB})$ e $\nu_0(\mathcal{BA})$ possono essere differenti, come mostra l'esempio

$$\mathcal{A} \leftarrow \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathcal{B} \leftarrow \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

con

$$\mathcal{AB} \leftarrow \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{BA} \leftarrow \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

n \neq ∞ Per operatori in spazi a infinite dimensioni può invece accadere che $\dim \ker(\mathcal{AB}) \neq \dim \ker(\mathcal{BA})$, ad es. si consideri $\mathcal{A} = \mathbf{a}^\dagger$, $\mathcal{B} = \mathbf{a}$, dove $\mathbf{a} \propto \mathbf{p} - i\mathbf{q}$ indica l'operatore di annichilazione in Meccanica Quantistica (si veda, per curiosità, Wikipedia alla voce "Index Theorem").



#21 Si veda [Gan59].

#24 L'identità è ovvia se si utilizza la decomposizione spettrale. Ma anche molto più direttamente si può calcolare così (ricordando che $\det(\mathbb{1} + \varepsilon \mathcal{X}) = 1 + \varepsilon \operatorname{Tr}(\mathcal{X}) + O(\varepsilon^2)$):

$$\det(e^{\mathcal{A}}) = \det(e^{\mathcal{A}/n})^n = \det(\mathbb{1} + \mathcal{A}/n + O(1/n^2))^n = (1 + \operatorname{Tr}(\mathcal{A})/n + O(1/n^2))^n \xrightarrow{n \rightarrow \infty} e^{\operatorname{Tr}(\mathcal{A})}.$$

#26 L'unitarietà discende dalle proprietà delle radici dell'unità:

$$\frac{1}{n+1} \sum_{j=0}^n \exp\left\{\frac{2\pi i}{n+1} j k\right\} = \begin{cases} 1 & \text{se } k = 0 \pmod{n+1} \\ 0 & \text{altrimenti} \end{cases}.$$

La matrice soddisfa l'identità $U^4 = \mathbb{1}$ con radici semplici $1, i, -1, -i$ e quindi lo spettro è contenuto in questo insieme di valori. Tuttavia determinare la rispettiva molteplicità non è un problema di semplice soluzione. Una soluzione brutale consiste nel chiedere a Mathematica, così

```
U[n_]:=Array[Exp[2 Pi I (#1-1) (#2-1)/n]/Sqrt[n] & , {n,n}];
(* .....Nota Bene.....^.....^.....^.....^..... *)
Eigenvalues[U[5]]
[-1, i, -i, 1, 1]
Eigenvalues[U[6]]
[-1, -1, i, -i, 1, 1]
```

etc, e tentare di estrarne una regolarità. La trattazione generale si trova su [Meh89], Sec.7.12 dove si mostra che la molteplicità è data da $r(1) = [(n+4)/4]$, $r(-1) = [(n+2)/4]$, $r(i) = [(n+1)/4]$, $r(-i) = [(n-1)/4]$, dove $[.]$ indica la "parte intera".

#31-32 Si veda [Lit50] Sec.6.2.

#34: Si veda [GL96], sec.2.1.3.

#38: Il modello è diffusamente descritto da Marc Kac in [Kac59, Wax54]. Ci limitiamo qui a pochi cenni. Si tratta di calcolare la potenza n -esima della matrice P_{jk} definita da

$$P_{jk} = \frac{1}{N} j \delta_{j+1, k} + \frac{1}{N} \delta_{j, k+1} (N - k + 1).$$

Dato che gli elementi di matrice sono funzioni polinomiali degli indici conviene associare al vettore \mathbf{p}_j il polinomio $p(x) = \sum_j p_j x^j$ e l'azione della matrice viene tradotta in quella di un operatore differenziale

$$P_{jk} p_k \implies (\mathcal{P} p)(x) = \left(x + \frac{1}{N}(1-x^2) \frac{d}{dx}\right) p(x)$$

Lo spettro di \mathcal{P} si determina facilmente attraverso l'equazione agli autovalori

$$\left(x + \frac{1}{N}(1-x^2)\frac{d}{dx}\right)p(x) = \lambda p(x)$$

Si trova $\lambda = 1 - 2n/N$, $n = 0, 1, 2, \dots, N$. Tutto lo spettro è contenuto nell'intervallo $[-1, 1]$ e per potenze grandi di \mathcal{P} domineranno gli autovalori 1 e -1 . L'autovettore appartenente all'autovalore 1 rappresenta la distribuzione di probabilità stazionaria cui ogni distribuzione tende per $t \rightarrow \infty$ (a patto di mediare sul tempo $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T p(t, :)$ cancellando così il termine oscillante $(-1)^t$).

#45 Se non si sa calcolare lo spettro significa che non si è letta l'introduzione – vedi Cap.1.3. In alternativa: si può considerare \mathcal{X} come un trancio finito dimensionale di una matrice infinita con elementi $X_{ij} = \delta_{i,j+1} + \delta_{j,i+1}$. Questa matrice ammette gli autovettori $x_j = (\dots, z^{-3}, z^{-2}, z^{-1}, 1, z, z^2, \dots)$ $y_j = (\dots, z^3, z^2, z, 1, z^{-1}, z^{-2}, \dots)$ con lo stesso autovalore $z + z^{-1}$. È sufficiente richiedere che il vettore $Ax_j + By_j$ si annulli per $j = 0$ e $j = N + 1$ per ottenere lo spettro.

#46 La relazione di ricorrenza permette di calcolare la riga $n + 1$ -esima dalla n -esima. Per costruzione la matrice commuta con \mathcal{X} del problema precedente e ne condivide perciò gli autovettori. Ne segue che

$$\sum_j \mathcal{Y}_{kj} \sin(\pi j n / (N+1)) = \lambda_n \sin(\pi k n / (N+1)), \quad (n = 1, 2, \dots, N)$$

La prima riga ci dà allora

$$\lambda_n = \sum_j \mathcal{Y}_{1j} \sin(\pi j n / (N+1)) / \sin(\pi n / (N+1))$$

Il codice `matlab` `hpm.m` costruisce la matrice a partire dalla prima riga. La parte centrale è data da

```

1      function Y=hpm(row)
2      ....
3      N=length(row);
4      U=eye(N,N);
5      X=diag(ones(N-1,1),-1);
6      X=X+X';
7      T=X;
8      Y=U*row(1);
9      Y=Y+X*row(2);
10     for n=3:N
11         Z=T*X-U;                                % relazione di ricorrenza
12         Y=Y+Z*row(n);
13         U=X;
14         X=Z;
15     end

```

Dalla linea 10 alla 15 si utilizza la relazione di ricorrenza del Probl. 48. Si potrebbe anche programmare direttamente $\mathcal{Y}_{n+1,m} = \mathcal{Y}_{n,m+1} + \mathcal{Y}_{n,m-1} - \mathcal{Y}_{n-1,m}$ ma questo richiede un doppio loop e un codice meno chiaro. Tuttavia (sorpresa!) il numero di operazioni richieste dalla regola di ricorrenza è molto superiore rispetto al codice “banale” e di conseguenza il tempo di esecuzione risulta anche superiore, il che non è rilevante per il calcolo una tantum, ma può essere rilevante per un codice utilizzato più volte come parte di un programma. Ad es. per dimensione 1000 i tempi sono rispettivamente 57 e 3 sec. Questo esempio dimostra che non sempre l'algoritmo più evoluto risulta più conveniente!

#47 La matrice è iperbolica, dunque la sua prima riga contiene tutta l'informazione necessaria, secondo quanto visto al Probl. 46. Può essere anche utilizzato un mezzo più diretto (*Mathematica*) ignorando questa proprietà ad es.

```

k1[i_, j_, n_] := Cot[Pi (i+j)/(2 (n+1))]^2 / (2 (n+1)) /; OddQ[i+j];
k1[i_, j_, n_] := 0 /; EvenQ[i+j];
k2[i_, j_, n_] := Cot[Pi (i-j)/(2 (n+1))]^2 / (2 (n+1)) /; OddQ[i+j];
k2[i_, j_, n_] := 0 /; EvenQ[i+j];
K[n_] := (n+1)/2 IdentityMatrix[n] +
Array[m1[#1, #2, n] - m2[#1, #2, n] &, {n, n}];
Eigenvalues[K[10]]/N]

```

(prudentemente si chiedi a *Mathematica* di calcolare in float a evitare un output gigantesco).

#48 La relazione di ricorrenza si dimostra controllando semplicemente la prima riga. Questa relazione definisce univocamente i “polinomi di Tschchebychev” di II tipo (si veda [GR65]).

#50 Si veda l'articolo [CMO80]. Lo spettro è dato da $m_j = \frac{1}{n+1} (nj - j^2/2)$. Si veda il codice *matlab/cmo.m*.

#51 Per costruzione la matrice è data dal polinomio

$$C = \sum_j C_{1,j} S^{j-1}$$

dove S è la matrice

$$S = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & \dots & \dots & 0 & 0 & 1 \\ 1 & 0 & 0 & \dots & \dots & 0 & 0 \end{pmatrix}$$

Ma $S^N = \mathbf{1}$ e dunque ...etc. Si veda [Meh89] per maggiori dettagli (oppure Wikipedia: circulant matrix).

#52: La relazione di commutazione dà

$$(\mathcal{JX})_{ij} = \delta_{i+1,k} X_{kj} = \begin{cases} X_{i+1,j} & \text{per } i < n \\ 0 & \text{per } i = n \end{cases}$$

$$(\mathcal{XJ})_{ij} = X_{ik} \delta_{k+1,j} = \begin{cases} X_{i,j-1} & \text{per } j > 1 \\ 0 & \text{per } j = 1 \end{cases}$$

ossia in modo più intuitivo

$$\mathcal{J} \mathcal{X} = \begin{pmatrix} X_{21} & X_{22} & X_{23} & \dots & X_{2n} \\ X_{31} & X_{32} & X_{33} & \dots & X_{3n} \\ X_{41} & X_{42} & X_{43} & \dots & X_{4n} \\ \vdots & \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & X_{n3} & \dots & X_{nn} \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

$$\mathcal{X} \mathcal{J} = \begin{pmatrix} 0 & X_{11} & X_{12} & \dots & X_{1n-1} \\ 0 & X_{21} & X_{22} & \dots & X_{2n-1} \\ 0 & X_{31} & X_{32} & \dots & X_{3n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & X_{n-11} & X_{n-12} & \dots & X_{n-1n-1} \\ 0 & X_{n1} & X_{n2} & \dots & X_{nn-1} \end{pmatrix}$$

il che implica che $X_{ij} = 0$ per $i > j$, $X_{11} = X_{22} = \dots = X_{nn}$ e in generale $X_{i+k} = x_k$, indipendente da i . Ne segue

$$\mathcal{X} = x_0 \mathbb{I} + x_1 \mathcal{J} + x_2 \mathcal{J}^2 + \dots + x_{n-1} \mathcal{J}^{n-1}$$

che è quanto si voleva dimostrare.

🔗 Sappiamo che se $[X, Y] = 0$ e se \mathcal{Y} ha tutti autovalori distinti allora $X = f(Y)$. Il risultato appena dimostrato è all'estremo opposto, tutti autovalori uguali, \mathcal{J} non è diagonalizzabile, ma la commutatività è limitata alle funzioni di \mathcal{J} . ◀

#56: L'equazione agli autovalori $Tx = \lambda x$ è semplicemente

$$\begin{aligned} t_{11} x_1 + t_{12} x_2 &= \lambda x_1 \\ t_{21} x_1 + t_{22} x_2 + t_{23} x_3 &= \lambda x_2 \\ t_{32} x_2 + t_{33} x_3 + t_{34} x_4 &= \lambda x_3 \\ &\vdots \\ t_{n-1n-2} x_{n-2} + t_{n-1n-1} x_{n-1} + t_{n-1n} x_n &= \lambda x_{n-1} \\ t_{nn-1} x_{n-1} + t_{nn} x_n &= \lambda x_n \end{aligned}$$

La prima equazione determina x_2 in funzione di x_1 ; la seconda determina x_3 in funzione di x_1, x_2 e via via fino ad arrivare a $(x_{n-2}, x_{n-1}) \rightarrow x_n$. L'ultima equazione determina λ ; l'unica costante arbitraria a disposizione rimane x_1 ; dunque per ogni autovalore esiste solo un autovettore. Se un elemento di matrice fuori diagonale si annulla la matrice si spezza nella somma diretta di due matrici tridiagonali nel qual caso si può avere degenerazione (autovalori multipli). Le matrici tridiagonali si prestano a una diagonalizzazione veloce attraverso l'algoritmo di Sturm (vedi [GL96]). L'algoritmo di Lanczos è basato sull'idea di trovare una trasformazione di similitudine che trasformi la matrice da diagonalizzare \mathcal{A} in una matrice tridiagonale

$$\mathcal{A} = \mathcal{S} \mathcal{T} \mathcal{S}^{-1}$$

essendo poi molto agevole trovare gli autovalori di \mathcal{T} . Si veda la sez. 4.5.3.

#58 Si veda [Hoc71]. Oppure si tenti una dimostrazione per induzione matematica partendo da $N = 2$.

#60 Dal problema precedente si sa che $p(x_0) = \langle e_{x_0} | p \rangle$. Dalla disuguaglianza di Cauchy si sa anche che

$$|p(x_0)| \leq \|e_{x_0}\| \|p\| = \|e_{x_0}\|$$

Dunque il polinomio cercato è $e_{x_0}(x) = K(x_0, x)$.

#63 Si tratta di un problema che ha legami con la teoria dei processi stocastici - lo stesso di un random walk sulle facce del cubo. Si può formalizzare il problema introducendo la matrice di adiacenza i cui indici sono in corrispondenza biunivoca con le facce e gli elementi di matrice sono uno se le facce sono adiacenti e zero altrimenti. Nel caso in esame la matrice di adiacenza è ovviamente

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

e l'operazione di media si realizza applicando $A/4$ al vettore $x = (1, 2, 3, 4, 5, 6)'$:

$$Ax = \begin{pmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 5/2 \\ 5/2 \end{pmatrix}$$

e dopo 16 volte

$$A^{16}x = \begin{pmatrix} 229375 \\ 229375 \\ 229375 \\ 229375 \\ 228378 \\ 229378 \end{pmatrix} / 65536$$

La teoria spettrale chiarisce il tutto. Gli autovalori di A sono $(1, -\frac{1}{2}, \frac{1}{4}, 0, 0, 0)$ e quindi in pochi passi il vettore x è proiettato sull'autovettore appartenente all'autovalore uno, il vettore con tutte componenti uguali. Dato che la somma dei valori sulle facce non cambia, il valore asintotico sarà $21/6$. La matrice A rappresenta un caso particolare di *matrice stocastica*.

#66 L'iterazione ha come soluzione

$$x_t = (\mathbb{1} - K)^{-1} (\mathbb{1} - K^t) y.$$

Se lo spettro di K è interamente contenuto nel disco $|\lambda| < 1$ allora $K^t \rightarrow \mathbb{O}$ per $t \rightarrow \infty$. Se invece anche un solo autovalore è in modulo maggiore o uguale a uno il procedimento non converge.

L'effetto di introdurre il parametro di rilassamento sostituisce a K l'operatore $(1 - \omega) + \omega K$ il cui spettro è ottenuto da quello di K attraverso una trasformazione che consiste in una contrazione verso il punto $z = 1$ (Fig.3-2).

$$(3.1) \quad 1 - z_\omega = \omega(1 - z).$$

La trasformazione può facilitare la convergenza "risucchiando" lo spettro di K dentro il disco unitario. A questo fine è necessario che gli autovalori di K giacciono a sinistra della retta $\Re\{z\} = 1$.

#69 È il problema di determinare lo spettro del momento angolare in meccanica quantistica. Le armoniche sferiche $r^\ell Y_\ell^m$ sono infatti una base per $\mathcal{H}_\ell^{(3)}$. Lo spettro di \mathcal{M}_3 è costituito da tutti gli interi m con $|m| \leq \ell$. Un calcolo di due righe è il seguente: introduciamo una coordinata

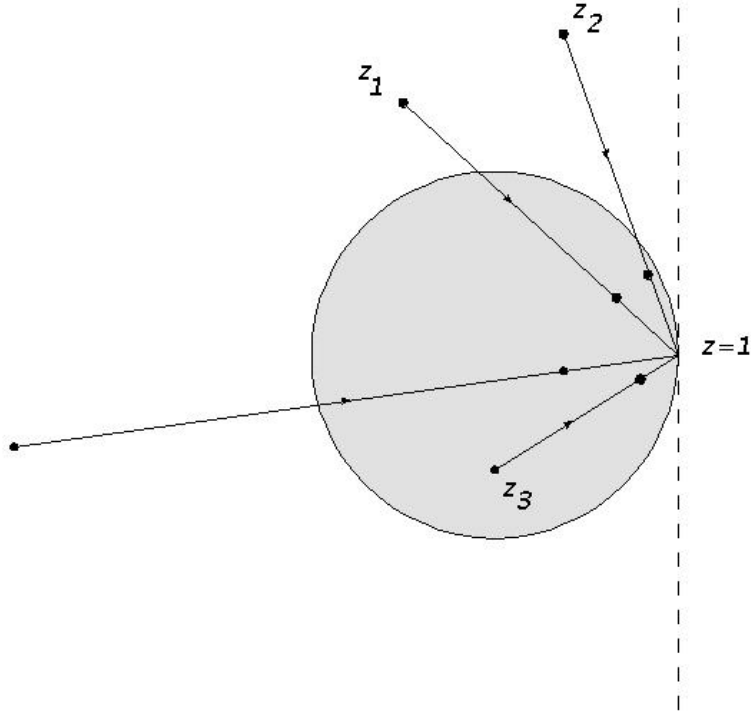


FIGURA 3-2. La trasformazione 3.1

complessa $\zeta = x + iy$. Allora si vede immediatamente che il generico polinomio omogeneo di grado ℓ è dato da

$$p(\zeta, \bar{\zeta}, z) = \sum_{0 \leq j, k \leq \ell} z^{\ell-j-k} \zeta^j \bar{\zeta}^k$$

e inoltre $\mathcal{M}_3 = \zeta \frac{\partial}{\partial \bar{\zeta}} - \bar{\zeta} \frac{\partial}{\partial \zeta}$, e quindi

$$(\mathcal{M}_3 - (j - k)) z^{\ell-j-k} \zeta^j \bar{\zeta}^k = 0$$

e lo spettro è dato dagli interi nell'intervallo $[-\ell, \ell]$. Un'analisi più accurata, che comprende la condizione che $p()$ sia armonico, permette di determinare le autofunzioni e la loro molteplicità (tutti autovalori semplici). Il nocciolo della dimostrazione è dato dal fatto che $r^\ell Y_\ell^\ell = \zeta^\ell$ è un polinomio armonico con autovalore $m = \ell$. Applicando l'operatore $\mathcal{M}_- = \mathcal{M}_1 - i\mathcal{M}_2$, dove $\mathcal{M}_i = i\varepsilon_{ijk} x_j \frac{\partial}{\partial x_k}$, si costruiscono tutti gli altri polinomi armonici (in numero $2\ell + 1$). Si può consultare un testo di M.Q. (ad es. [OD96]) oppure [Ham62, Ste94].

#73 Se \mathcal{A} è diagonale ($\mathcal{A} = a_i \delta_{ij}$), si ha per ogni matrice

$$[\mathcal{A}, \mathcal{B}]_{ij} = (a_i - a_j) B_{ij}$$

e pertanto

$$[\mathcal{A}, [\mathcal{A}, \mathcal{B}]]_{ij} = (a_i - a_j)^2 B_{ij} = 0$$

implica $(a_i - a_j) B_{ij} = Z_{ij} = 0$.

#74 Per ipotesi $\det(z\mathbb{1} - A - tB) = D(z)$. Raccogliamo t per ottenere $\det((z\mathbb{1} - A)/t - B) = D(z)/t^n$. Posto $z = t\zeta$, si ha $\det((\zeta\mathbb{1} - A/t - B) = D(t\zeta)/t^n$ e al limite per $t \rightarrow \infty$ si ottiene $\det(\zeta\mathbb{1} - B) = \zeta^n$, da cui $B^n = \mathbb{O}$.

#76 Si prenda la traccia: $\text{Tr}(\mathbf{AB}) - \text{Tr}(\mathbf{BA}) = 0$ per la proprietà ciclica della traccia, mentre $\text{Tr}(\mathbf{1}) = \dim \mathbf{1}$.

n ≠ ∞ In infinite dimensioni c'è la possibilità di aggirare questo ostacolo, se non altro il concetto di traccia non si applica a tutti gli operatori, e infatti la regola di commutazione $[\mathcal{A}, \mathcal{B}] = \mathbf{1}$ è uno dei mattoni fondamentali della Meccanica Quantistica. ◀

#78 L'operatore \mathcal{A} si può rappresentare come il prodotto di due matrici

$$\begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_r & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} \mathbf{b}'_1 \\ \mathbf{b}'_2 \\ \vdots \\ \mathbf{b}'_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

dove \mathbf{a}_i e \mathbf{b}_i sono vettori colonna e \mathbf{b}'_i (trasposto) sono dei vettori riga. Ma sappiamo dal Probl. 20 che possiamo permutare le due matrici nel calcolo del polinomio caratteristico, ne segue immediatamente il risultato. Alternativamente, si può argomentare come segue. L'operatore $\mathcal{X} = \sum_j |\mathbf{a}_j\rangle \langle \mathbf{b}_j|$ soddisfa $\mathcal{X}^2 = \sum_{jk} |\mathbf{a}_j\rangle \langle \mathbf{b}_j| \mathbf{a}_k \rangle \langle \mathbf{b}_k|$; indichiamo con \mathcal{X}_s la matrice $\langle \mathbf{b}_i | \mathbf{a}_j \rangle$. Allora

$$\mathcal{X}^{n+1} = \sum_{ij} |\mathbf{a}_i\rangle (\mathcal{X}_s^n)_{ij} \langle \mathbf{b}_j|$$

Se $\Delta_s(z)$ è il polinomio caratteristico di \mathcal{X}_s ne segue che $\Delta(\mathcal{X}) = 0$ con $\Delta(z) = z \Delta_s(z)$, il che porta alla stessa conclusione.

#80 È un caso speciale del problema 78. Infatti, definiamo la matrice come segue:

$$\mathcal{A} \leftrightarrow \begin{pmatrix} \mathbf{a}_1 + \mathbf{c}_1 & \mathbf{b}_1 + \mathbf{c}_2 & \mathbf{c}_3 & \dots & \mathbf{c}_n \\ \mathbf{a}_2 + \mathbf{d}_1 & \mathbf{b}_2 + \mathbf{d}_2 & \mathbf{d}_3 & \dots & \mathbf{d}_n \\ \mathbf{a}_3 & \mathbf{b}_3 & 0 & \dots & 0 \\ \mathbf{a}_4 & \mathbf{b}_4 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \dots & \dots \\ \mathbf{a}_n & \mathbf{b}_n & 0 & \dots & 0 \end{pmatrix}$$

L'azione di \mathcal{A} su un vettore $\mathbf{x} \leftrightarrow (x_1, \dots, x_n)'$ dà

$$x_1 \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} + x_2 \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} + \langle \mathbf{c} | \mathbf{x} \rangle \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \langle \mathbf{d} | \mathbf{x} \rangle \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

ossia in definitiva

$$\mathcal{A} = |\mathbf{a}\rangle \langle \mathbf{1}| + |\mathbf{b}\rangle \langle \mathbf{2}| + |\mathbf{c}\rangle \langle \mathbf{1}| + |\mathbf{d}\rangle \langle \mathbf{2}|$$

e perciò si riduce il problema al calcolo dello spettro di una matrice 4×4

$$\begin{pmatrix} \mathbf{a}_1 & \mathbf{b}_1 & 1 & 0 \\ \mathbf{a}_2 & \mathbf{b}_2 & 0 & 1 \\ \langle \mathbf{c} | \mathbf{a} \rangle & \langle \mathbf{c} | \mathbf{b} \rangle & \mathbf{c}_1 & \mathbf{c}_2 \\ \langle \mathbf{d} | \mathbf{a} \rangle & \langle \mathbf{d} | \mathbf{b} \rangle & \mathbf{d}_1 & \mathbf{d}_2 \end{pmatrix}$$

Il vantaggio, dal punto di vista computazionale, è evidente nel caso in cui $n \gg 4$. Si può facilmente verificare la cosa con *matlab* scegliendo ad es. $n \gtrsim 1000$. Si nota che lo spettro

dipende solo dalle componenti di indice 1,2 e dai prodotti scalari dei vettori $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$. Il fatto si può giustificare considerando la trasformazione di similitudine

$$\mathcal{A} \rightarrow \mathcal{R} \mathcal{A} \mathcal{R}'$$

con

$$\mathcal{R} \leftrightarrow \left(\begin{array}{c|c} \mathbb{1}_{2,2} & \mathbb{O}_{2,n} \\ \hline \mathbb{O}_{n,2} & \mathcal{R}^{(n-2)} \end{array} \right)$$

corrisponde a una rotazione nel sottospazio $(3, 4, \dots, n)$ se $\mathcal{R}^{(n-2)}$ è una matrice ortogonale di dimensione $n-2$.

#81 Il risultato è facilmente verificabile in pochi secondi utilizzando un programma simbolico come Mathematica.

```
H0[n_]:=SparseArray[{i_,j_}-> i,{n,n}];
V[n_]:=SparseArray[{i_,j_}/;Abs[i-j]=1->1,{n,n}];
H[n_]:=H0[n]+x V[n];
Series[Eigenvalues[H[20],{x,0,14}]
```

Tuttavia una dimostrazione formale non è difficile da ottenere sfruttando un po' di algebra. Sia

$$\mathcal{W}_{ij} = \delta_{i,j+1} - \delta_{j,i+1}.$$

\mathcal{W} è antisimmetrica, per cui $\exp\{\lambda \mathcal{W}\}$ è unitaria. Si trovano subito le proprietà di commutazione

$$[\mathcal{W}, \mathcal{H}_0] = -\mathcal{V}, \quad [\mathcal{W}, \mathcal{V}] = -2|1\rangle\langle 1| + 2|N\rangle\langle N|$$

Si ha allora, applicando la Eq.3.13

$$\begin{aligned} e^{\lambda \mathcal{W}} (\mathcal{H}_0 + \lambda \mathcal{V}) e^{-\lambda \mathcal{W}} &= \mathcal{H}_0 - \lambda \mathcal{V} + \sum_{k \geq 2} \frac{\lambda^k}{k!} \text{ad}_{\mathcal{W}}^k (\mathcal{H}_0) + \lambda \mathcal{V} + \sum_{k \geq 1} \frac{\lambda^k}{k!} \text{ad}_{\mathcal{W}}^k (\mathcal{V}) \\ &= \mathcal{H}_0 + \lambda \tilde{\mathcal{V}} \end{aligned}$$

dove si utilizza la notazione usuale $\text{ad}_X(Y) \equiv [X, Y]$. Ciò mostra che l'operatore $\mathcal{H}_0 + \lambda \mathcal{V}$ è unitariamente equivalente a $\mathcal{H}_0 + \lambda \tilde{\mathcal{V}}$ e si possono calcolare gli elementi di matrice di $\tilde{\mathcal{V}}$ come segue:

$$\begin{aligned} \lambda \tilde{\mathcal{V}} &= \sum_{k \geq 0} \frac{\lambda^{k+2}}{(k+2)!} \text{ad}_{\mathcal{W}}^{k+2} (\mathcal{H}_0) + \lambda \sum_{k \geq 0} \frac{\lambda^{k+1}}{(k+1)!} \text{ad}_{\mathcal{W}}^{k+1} (\mathcal{V}) \\ &= -2 \sum_{k \geq 0} \frac{(k+1) \lambda^{k+2}}{(k+2)!} \text{ad}_{\mathcal{W}}^k (\mathbb{P}_1 - \mathbb{P}_N) \end{aligned}$$

L'elemento di matrice diagonale di $\langle n | \tilde{\mathcal{V}} | n \rangle$, che ci dà la correzione principale all'autovalore n -esimo si può calcolare tenendo conto che si ha $\langle n | \mathcal{W}^m | 1 \rangle = 0$ per $m < n-1$ e quindi il primo contributo si ha per $k = 2n-2$, cosicché la correzione è proporzionale a λ^{2n} . Il coefficiente di $\mathcal{W}^{n-1} \mathbb{P}_1 \mathcal{W}^{n-1}$ nel commutatore multiplo $\text{ad}_{\mathcal{W}}^{2n-2} \mathbb{P}_1$, si trova immediatamente utilizzando lo sviluppo dell'Es.62: si trova $(-1)^{n-1} \binom{2n-2}{n-1}$ da cui $\delta_{2n} = \frac{1}{(n-1)!n!}$, in quanto il segno è assorbito dal segno di $\langle 1 | \mathcal{W}^{n-1} | n \rangle$. Quando $n > N/2$ diventa dominante l'elemento di matrice $\langle N | \mathcal{W}^k | N \rangle$; la correzione risulta anti-simmetrica rispetto a $N/2$ e la somma di tutti le correzioni è nulla (come deve essere per via che $\text{Tr}(\mathcal{V}) = 0$).

III - Operatori lineari nello spazio di Hilbert

CAPITOLO 4

Spazio di Hilbert

4.1. Spazi lineari ad infinite dimensioni

4.1.1. Definizioni fondamentali. Sia \mathcal{H} uno spazio lineare complesso dotato di prodotto interno $\langle \cdot | \cdot \rangle$. Supporremo in questa terza parte delle lezioni che lo spazio abbia dimensione infinita, e cioè esista un numero n di vettori linearmente indipendenti per ogni $n \in \mathbb{Z}^+$. Ecco alcuni esempi.

- i) ℓ_2 - lo spazio di tutte le successioni $(z_1, z_2, \dots, z_n, \dots \in \mathbb{C})$ tali che la somma $\sum_{n=1}^{\infty} |z_n|^2$ sia convergente; il prodotto interno è dato da

$$\langle z | w \rangle \equiv \sum_{n=1}^{\infty} \overline{z_n} w_n.$$

- ii) $C_2(a, b)$ - lo spazio di tutte le funzioni continue nell'intervallo reale $(a, b) \mapsto \mathbb{C}$ tali che esista finito l'integrale

$$\int_a^b |f(x)|^2 dx$$

il prodotto interno essendo dato da

$$\langle f | g \rangle \equiv \int_a^b \overline{f(x)} g(x) dx.$$

- iii) $H_2(\mathbb{D})$ - lo spazio di tutte le funzioni analitiche regolari in una regione limitata \mathbb{D} nel piano complesso per le quali esista finito l'integrale

$$\int_{\mathbb{D}} |f(z)|^2 [dz]$$

dove* $[dz] = \overline{dz} \wedge dz / 2i$; il prodotto interno è dato da

$$\langle f | g \rangle \equiv \int_{\mathbb{D}} \overline{f(z)} g(z) [dz].$$

◀ Avremo modo di considerare altri esempi di spazi a ∞ -dimensioni; un ruolo fondamentale spetta agli spazi $\mathcal{L}_2(X, d\mu(x))$, dove X è un sottoinsieme di \mathbb{R}^n e $d\mu(x)$ una misura di integrazione. La definizione più conveniente di \mathcal{L}_2 è in termini di “funzioni integrabili secondo Lebesgue”. Nel seguito tuttavia eviteremo l'uso di tale nozione e definiremo \mathcal{L}_2 come “complemento” di C_2 . ▶

In uno spazio ad infinite dimensioni valgono tutte le proprietà già dimostrate nella Parte II a patto che siano coinvolti solo vettori in numero finito. Ad esempio valgono le disuguaglianze considerate nell'App. 3.6 a p. 76.

Dati n vettori ortonormali (e_1, e_2, \dots, e_n) è unica la decomposizione del vettore $\xi \in \mathcal{H}$

$$\xi = \sum_{i=1}^n \xi_i e_i + \xi_{\perp}$$

*La notazione $\overline{dz} \wedge dz$ è quella delle forme differenziali; equivale all'elemento d'area $2i dx dy$.

con $\langle \xi_\perp | e_i \rangle = 0$. Inoltre

$$\langle \xi | \xi \rangle \geq \sum_{i=1}^n |\langle e_i | \xi \rangle|^2$$

che viene chiamata “disuguaglianza di Bessel”.

Una successione di vettori $\{\xi_n | n = 1, 2, \dots, \infty\}$ si dice *convergente* con limite ξ se $\|\xi - \xi_n\| \rightarrow 0$ per $n \rightarrow \infty$. Dato un insieme \mathcal{S} di vettori in \mathcal{H} si indica con $\overline{\mathcal{S}}$ e si chiama *chiusura* di \mathcal{S} l'insieme formato da tutti i punti limite di \mathcal{S} , cioè tutti i vettori ξ per cui esista una successione avente ξ come limite. \mathcal{S} si dice *chiuso* se $\overline{\mathcal{S}} \equiv \mathcal{S}$. \mathcal{S} si dice *denso* in \mathcal{H} se $\overline{\mathcal{S}} = \mathcal{H}$. Una successione si dice *di Cauchy* se per ogni $\varepsilon > 0$ prefissato esiste un N tale che

$$\|\xi_j - \xi_k\| < \varepsilon \quad \forall (j, k) > N$$

Si dimostra facilmente che ogni successione convergente è “di Cauchy”, *ma non vale in generale il contrario*.

Def.4.1.1

DEF. 4.1.1. Uno spazio lineare dotato di norma $\|\cdot\|$ si dice *completo* se ogni successione di Cauchy converge ad un vettore dello spazio.

Ogni spazio a dimensione finita risulta completo, ma ciò non è vero per spazi ad infinite dimensioni. Lo spazio C_2 ad es. non è completo. Per dimostrarlo è sufficiente esibire anche una sola successione di Cauchy non-convergente.

ESEMPIO 4-1. Definiamo la successione in $C_2(0, 1)$ come segue, indicando con x_n il punto $x_n = 1/2 - 1/n$:

$$f_n(x) = \begin{cases} 0 & \text{per } 0 \leq x < x_n \\ \frac{x - x_n}{1/2 - x_n} & \text{per } x_n \leq x < 1/2 \\ 1 & \text{per } 1/2 \leq x \leq 1 \end{cases}$$

Si ha (per $n < m$)

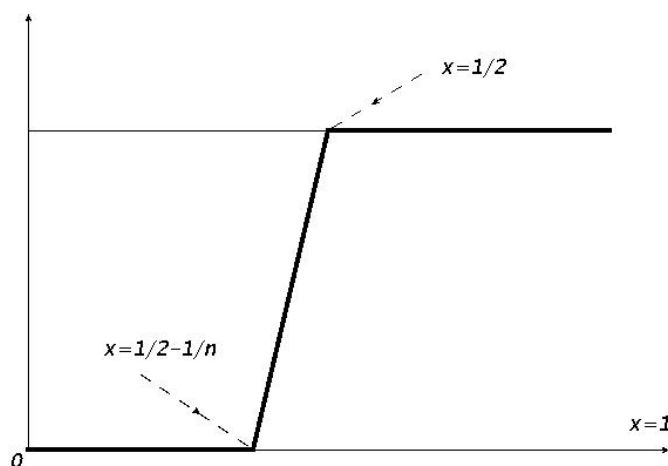


FIGURA 4-1. La successione di Cauchy non convergente in C_2

$$\|f_n - f_m\|^2 = \frac{2}{3} \frac{(x_n - x_m)^2}{1 - 2x_n} < \frac{1}{3} (1/2 - x_n) = \frac{1}{3n}$$

dunque la successione è di Cauchy. D'altra parte non esiste alcuna funzione continua $f(x)$ tale che $\lim \|f_n - f\| = 0$. ◀

Un criterio per decidere se uno spazio è completo è il seguente:

TEOREMA 4.1.1. *Uno spazio metrico \mathcal{H} è completo se e solo se ogni serie assolutamente convergente è convergente.*

Teor.4.1.1

⚡ Non si tratta di un gioco di parole: una serie $\sum_n \xi_n$ è assolutamente convergente se converge la serie $\sum_n \|\xi_n\|$. ◀

PROVA. Sia \mathcal{H} completo e assumiamo che $\sum_{n=1}^\infty \xi_n$ sia assolutamente convergente. Da $\sum_{n=1}^\infty \|\xi_n\| = M$ segue che per ogni $\varepsilon > 0$ esiste un N_ε tale che

$$\sum_{n \geq N_\varepsilon} \|\xi_n\| < \varepsilon$$

ne segue che la successione $\eta_n = \sum_{j=1}^n \xi_j$ è di Cauchy e dunque convergente. Viceversa assumiamo che ogni successione assolutamente convergente sia convergente. Sia η_n una successione di Cauchy: esiste una sottosuccessione $\{\eta_{j_1}, \eta_{j_2}, \dots, \eta_{j_k}, \dots\}$ tale che

$$\|\eta_{j_n} - \eta_{j_m}\| < 2^{-k}$$

per $n > m > k$. La serie

$$\eta_{j_1} + \sum_{k=1}^\infty (\eta_{j_{k+1}} - \eta_{j_k})$$

risulta assolutamente convergente e secondo l'ipotesi converge a un vettore $\xi \in \mathcal{H}$. Si ha d'altronde

$$\|\xi - \eta_n\| = \lim_{k \rightarrow \infty} \|\eta_{j_k} - \eta_n\| < \varepsilon \quad (n > N)$$

■

TEOREMA 4.1.2. ℓ_2 è completo.

Teor.4.1.2

PROVA. Sia $(\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(n)}, \dots) \in \ell^2$ con $\|\xi^{(n)} - \xi^{(m)}\| < \varepsilon$ ($\forall n > m > N_\varepsilon$). Si ha ovviamente che le successioni delle componenti dei vettori sono di Cauchy, in quanto

$$|\xi_j^{(n)} - \xi_j^{(m)}| \leq \|\xi^{(n)} - \xi^{(m)}\|$$

e perciò esiste il limite, per ogni j , $\lim_n \xi_j^{(n)} = \xi_j$. Resta da dimostrare che $\sum_n |\xi_n|^2$ è convergente. La catena di implicazioni seguente dimostra la tesi:

$$\begin{aligned} \sum_{j=1}^M |\xi_j^{(n)} - \xi_j^{(m)}|^2 &< \varepsilon^2 \\ &\downarrow \lim_{m \rightarrow \infty} \\ \sum_{j=1}^M |\xi_j^{(n)} - \xi_j|^2 &< \varepsilon^2 \\ &\downarrow \lim_{M \rightarrow \infty} \\ \sum_{j=1}^\infty |\xi_j^{(n)} - \xi_j|^2 &< \varepsilon^2 \\ &\downarrow \\ \sum_{j=1}^M |\xi_j|^2 = \sum_{j=1}^M |\xi_j^{(n)} - \xi_j + \xi_j^{(n)}|^2 &\leq \left(\sqrt{\sum_{j=1}^M |\xi_j^{(n)} - \xi_j|^2} + \sqrt{\sum_{j=1}^M |\xi_j^{(n)}|^2} \right)^2 \end{aligned}$$

ed il limite esiste quindi finito. ■

⚡ Da notare che la condizione essenziale che permette di raggiungere la dimostrazione è che ogni componente ξ_j di $\xi \in \ell_2$ soddisfa la disuguaglianza $|\xi_j| \leq \|\xi\|$. In generale si parla di “spazi

funzionali” per gli spazi ad infinite dimensioni i cui elementi sono funzioni da un insieme $\mathcal{M} \mapsto \mathbb{C}$ tali che

$$|f(\xi)| \leq K(\xi) \|f\|$$

Lo spazio $\mathfrak{H}_2(\mathbb{D})$ ne è un esempio essendo

$$|f(z)| \leq K(\bar{z}, z) \|f\|$$

(vedi Probl. 60). La funzione $K(\bar{w}, z)$ è chiamata il nucleo di Bergman del dominio \mathbb{D} . Lo spazio \mathfrak{H}_2 è completo (si veda [Hel78], Ch.VIII). ◀

🔗 In Fisica Teorica gli spazi di Hilbert costituite da funzioni analitiche in vari domini complessi a più dimensioni, in particolare lo spazio di Bargmann, si presentano in modo naturale nello studio degli stati coerenti e nella teoria della quantizzazione geometrica. Per un ampio panorama sull’argomento si può consultare la raccolta di Klauder–Skagerstam [KS85]. ◀

Ogni spazio non-completo può essere completato, cioè può essere realizzato come sottospazio denso di uno spazio completo. Il procedimento è analogo a quello che definisce il campo dei numeri reali a partire dai razionali. Dato lo spazio \mathcal{H} (non completo) si costruisce lo spazio $\overline{\mathcal{H}}$ i cui elementi sono le successioni di Cauchy in \mathcal{H} :

$$\overline{\mathcal{H}} = \{\Xi \mid \Xi = (\xi_1, \xi_2, \dots, \xi_n, \dots) \in \mathcal{H}\}$$

e si pone

$$\langle \Xi \mid \Xi' \rangle \equiv \lim_{n \rightarrow \infty} \langle \xi_n \mid \xi'_n \rangle$$

(il limite esiste in quanto $\langle \xi_n \mid \xi'_n \rangle$ è una successione numerica di Cauchy). Il passo successivo consiste nell’identificare i vettori Ξ e Ξ' se

$$\|\Xi - \Xi'\| \equiv \lim_{n \rightarrow \infty} \|\xi_n - \xi'_n\| = 0$$

Si scriverà allora $\Xi \sim \Xi'$ e lo spazio delle classi di equivalenza $\mathfrak{H} = \overline{\mathcal{H}} / \sim$ è uno spazio completo, che contiene \mathcal{H} come sottospazio denso.

Def.4.1.2

DEF. 4.1.2. Uno spazio metrico completo \mathfrak{H} è detto uno *spazio di Hilbert*.

Nelle applicazioni si considerano per lo più spazi di Hilbert che soddisfano la condizione supplementare di *separabilità*.

Def.4.1.3

DEF. 4.1.3. Uno spazio di Hilbert \mathfrak{H} si dice *separabile* se esso contiene una successione $\xi_1, \xi_2, \dots, \xi_n, \dots$ ovunque densa. Cioè per ogni $\varepsilon > 0$ e per ogni vettore $\xi \in \mathcal{H}$ esiste un k_ε tale che

$$\|\xi - \xi_{k_\varepsilon}\| < \varepsilon$$

🔗 La successione cui si fa riferimento nella definizione di separabilità è l’analogo della successione dei numeri razionali che è ovunque densa nel campo reale. ◀

Ci occuperemo solo di spazi di Hilbert separabili, per cui ometteremo il termine “separabile” a meno che non sia opportuno sottolinearlo. Si può dimostrare che ogni spazio di Hilbert separabile ammette basi ortonormali complete, ed anzi ogni base ortonormale è numerabile. Una base ortonormale è data da un insieme di vettori

$$e_1, e_2, \dots, e_n, \dots$$

con $\langle e_i \mid e_j \rangle = \delta_{ij}$, con la proprietà che

$$\forall j : \langle e_j \mid \xi \rangle = 0 \implies \xi = 0.$$

È chiaro che non è sufficiente costruire un insieme infinito ortonormale per avere una base; l'ultima condizione è essenziale per fare sì che $\xi = \sum_j |e_j\rangle \langle e_j| \xi$, o equivalentemente

$$\sum_j |e_j\rangle \langle e_j| = \mathbb{1}$$

detta condizione di completezza dell'insieme ortonormale $\{e_n\}$. ◀

ESEMPIO 4-2 (ℓ_2). Abbiamo già mostrato che ℓ_2 è completo; inoltre esso è evidentemente separabile, dato che ammette la base ortonormale completa

$$\{e_j = (\underbrace{0, 0, \dots, 0}_{\leftarrow j \rightarrow}, 1, 0, \dots) \mid j = 1, 2, \dots, \infty\}$$

ℓ_2 è il modello fondamentale di spazio di Hilbert separabile; ogni altro spazio siffatto può essere messo in corrispondenza biunivoca (e isometrica) con ℓ_2 mediante l'introduzione di una base ortonormale:

$$\xi = \sum \xi_i e_i \in \mathfrak{H} \mapsto (\xi_1, \xi_2, \dots, \xi_n, \dots) \in \ell_2.$$

ESEMPIO 4-3 (L_2). Consideriamo lo spazio $C_2(a, b)$ delle funzioni continue definite nell'intervallo (a, b) . Sappiamo che lo spazio non è completo. Il procedimento generale di completamento porta allo spazio $\mathcal{L}_2(a, b)$. Questo spazio non è costituito da un insieme di funzioni, bensì da classi di equivalenza di successioni di Cauchy di funzioni continue; questo fatto rende piuttosto complicata la geometria di questo spazio. È equivalente definire \mathcal{L}_2 come lo spazio delle funzioni “misurabili secondo Lebesgue”, ma anche secondo questa definizione è necessario considerare le classi di equivalenza rispetto alla relazione $f \sim g$ se $f - g = 0$ quasi dappertutto. Nelle applicazioni, tuttavia, sono alcuni sottospazi di \mathcal{L}_2 , ad es. quello delle funzioni assolutamente continue, che entreranno in gioco nella teoria degli operatori lineari. \mathcal{L}_2 è completo e separabile. Una base ortonormale numerabile è fornita ad esempio dai polinomi ortogonali rispetto ad una misura $\int_a^b \mu(dx) = 1$, come introdotti nell'appendice 3.5.

ESEMPIO 4-4 (H_2). $H_2(\mathbb{D})$ è completo e separabile. Ad es. per $\mathbb{D} \equiv \{z \mid |z| < 1\}$ una base ortonormale è data da

$$e_n(z) = \sqrt{\frac{n+1}{\pi}} z^n, \quad (n = 0, 1, 2, \dots, \infty).$$

Un generico vettore di $H_2(\mathbb{D})$ è dato da una serie di potenze del tipo

$$f(z) = \sum_{n=0}^{\infty} f_n \sqrt{\frac{n+1}{\pi}} z^n$$

dove la serie $\sum |f_n|^2$ è convergente. Per un dominio complesso differente, ad es. $\mathbb{D} = \{z \in \mathbb{C} \mid 1 < |z| < 2\}$ una base è fornita da tutte le potenze (positive e negative) e il corrispondente sviluppo sui vettori di base equivale allo sviluppo di Laurent. In ogni caso $H_2(\mathbb{D})$ risulta essere un sottospazio chiuso di $\mathcal{L}_2(\mathbb{D})$ e di conseguenza completo e separabile [Hel78].

ESEMPIO 4-5. La somma diretta $\sum_{n \geq 1}^{\oplus} \mathfrak{H}_n$ di spazi di Hilbert, definita come lo spazio delle N-ple ordinate $\{f_n \in \mathfrak{H}_n \mid n = 1, 2, \dots, N\}$ con la norma definita da $\|f\|^2 = \sum_{n=1}^N \|f_n\|^2$ costituisce uno spazio di Hilbert separabile (lo spazio di Fock in fisica quantistica costituisce un esempio con $N \rightarrow \infty$). ◀

4.1.2. Funzionali lineari. Un funzionale lineare nello spazio di Hilbert \mathfrak{H} è definito come nel caso finito-dimensionale. Tuttavia ci si rende conto immediatamente che non tutti i funzionali lineari possono essere definiti dappertutto in \mathfrak{H} (a meno di non operare una definizione convenzionale impiegando le cosiddette *basi di Hamel* [Mad70, Tay58]). Se ad es. $z \in \ell_2 = (z_1, z_2, \dots, z_n, \dots)$, definiamo

$$\phi(z) = \sum_{j=1}^{\infty} z_j.$$

È chiaro che esistono vettori in ℓ_2 per cui la serie che definisce ϕ diverge (ad. es. $z_j = 1/j$). Un funzionale lineare sarà perciò definito in un opportuno sottospazio lineare D_ϕ detto il *dominio* di ϕ

$$\phi \in D_\phi \subset \mathfrak{H} \mapsto \mathbb{C}.$$

Def.4.1.4

DEF. 4.1.4. Un funzionale lineare ϕ con dominio D_ϕ è detto *limitato* se esiste un numero positivo M tale che

$$|\phi(\xi)| \leq M \|\xi\| \quad (\forall \xi \in D_\phi).$$

Teor.4.1.3

TEOREMA 4.1.3. *Un funzionale lineare (ϕ, D_ϕ) è limitato se e solo se esso è continuo.*

PROVA. Sia $\{\xi_n | n = 1, 2, \dots\}$ una successione di vettori in D_ϕ avente limite $\xi \in D_\phi$. Da $\|\xi - \xi_n\| \rightarrow 0$ segue anche

$$|\phi(\xi_n) - \phi(\xi)| = |\phi(\xi - \xi_n)| \leq M \|\xi_n - \xi\| \rightarrow 0$$

dunque ϕ è continuo. Viceversa, se ϕ è continuo segue che ϕ è limitato: infatti supponiamo che ϕ non sia limitato, e perciò che per ogni n esista un vettore ξ_n tale che $|\phi(\xi_n)| > n \|\xi_n\|$. Ne segue

$$\phi\left(\frac{\xi_n}{n \|\xi_n\|}\right) > 1$$

ma la successione $\frac{\xi_n}{n \|\xi_n\|}$ tende a zero il che contraddice la continuità. ■

Def.4.1.5

DEF. 4.1.5. Un funzionale lineare si dice *densamente definito* se la chiusura del suo dominio è tutto lo spazio $\overline{D_\phi} = \mathfrak{H}$.

Un funzionale lineare densamente definito si può estendere per continuità a tutto lo spazio conservando le sue proprietà.

ESEMPIO 4-6. Sia ad es. $\mathfrak{H} = \mathcal{L}_2(0, 1)$ e

$$\phi(f) = \int_0^1 f(x) dx, \quad D_\phi = \{f \in C_2(0, 1) \subset \mathcal{L}_2(0, 1)\}$$

cioè il dominio è costituito dalle funzioni continue a quadrato sommabile. Dalla disuguaglianza[†]

$$\left| \int_0^1 f(x) dx \right| \leq \sqrt{\int_0^1 |f(x)|^2 dx}$$

segue che il funzionale è limitato.

Il teorema che segue rappresenta l'analogo di quello dimostrato nella Parte II (2.4.2 a p. 31):

RIESZ

TEOREMA 4.1.4 (di Riesz). *Il più generale funzionale lineare limitato è della forma*

$$\phi(\xi) = \langle \phi^\sharp | \xi \rangle.$$

[†]Conseguenza di $\int_0^1 |f(x) - f(y)|^2 dx \geq 0$.

PROVA. Diamo due dimostrazioni, una elementare e una indipendente dalla introduzione di una specifica base. Siano $\{e_1, e_2, \dots, e_n, \dots\}$ i vettori di una base ortonormale. Sia $\xi = \sum \xi_i e_i$; per la continuità di ϕ si avrà

$$\phi(\xi) = \sum_i \xi_i \phi(e_i)$$

il che suggerisce di porre $\phi^\sharp = \sum_i \overline{\phi(e_i)} e_i$; si tratta di dimostrare che $\phi^\sharp \in \mathfrak{H}$. Sia $\phi^{(N)} = \sum_{i=1}^N \overline{\phi(e_i)} e_i$; si trova $\phi(\phi^{(N)}) = \|\phi^{(N)}\|^2$ e dalla limitatezza del funzionale

$$\|\phi^{(N)}\|^2 \leq M \|\phi^{(N)}\|$$

segue $\|\phi^{(N)}\| \leq M$. Il limite $N \rightarrow \infty$ definisce perciò un vettore in \mathfrak{H} . 2^a dimostrazione: sia $\ker(\phi)$ il sottospazio costituito dai vettori ξ tali che $\phi(\xi) = 0$; il punto cruciale è che la continuità di ϕ comporta che $\ker(\phi) \subset \mathfrak{H}$ è un sottospazio chiuso e se ϕ non è identicamente nullo $\ker(\phi)$ è un sottospazio proprio di \mathfrak{H} . Dunque esiste un vettore η ortogonale a $\ker(\phi)$ e di più questo vettore è unico a meno di un fattore moltiplicativo. Si ha perciò

$$\xi = \xi_0 + \alpha \eta, \quad \phi(\xi_0) = 0, \quad \langle \eta, \xi_0 \rangle = 0$$

Ne segue

$$\phi(\xi) = \alpha \phi(\eta), \quad \langle \eta, \xi \rangle = \alpha \|\eta\|^2$$

Eliminando α si ottiene

$$\phi(\xi) = \frac{\phi(\eta)}{\|\eta\|^2} \langle \eta, \xi \rangle$$

e infine

$$\boxed{\phi^\sharp = \frac{\overline{\phi(\eta)} \eta}{\|\eta\|^2}}$$

Si verifica così che ϕ^\sharp dipende solo dalla direzione di η e cioè da $\ker(\phi)$. ■

In conclusione lo spazio dei funzionali lineari continui su uno spazio di Hilbert è identificabile con lo spazio stesso in modo del tutto naturale, ed è consistente utilizzare la notazione di Dirac anche per lo spazio di Hilbert. Nell'esempio fornito in precedenza (Es. 4-6) in effetti $\phi(f)$ coincide con il prodotto interno di f con la funzione identicamente uguale a uno. In $L_2(0, \infty)$ invece la stessa espressione non definisce un funzionale continuo. Ci si può chiedere che cosa venga a mancare, nella derivazione precedente, nel caso di un funzionale lineare non-limitato. Una possibilità è che non esista alcun vettore ortogonale a $\ker(\phi)$, ossia che $\ker \phi$ sia denso in D_ϕ (come nell'esempio che segue).

ESEMPIO 4-7. Un esempio importante di funzionale lineare è il seguente. Sia $\mathfrak{H} = L_2(-1, 1)$ e $\delta(f) = f(0)$, $D_\delta = C_2(-1, 1)$. Il funzionale lineare così definito risulta essere illimitato e perciò sfugge al teorema di Riesz: non esiste alcun vettore in L_2 tale che sia $\delta(f) = \langle \delta | f \rangle$, anche se nella pratica corrente si scrive spesso $\int \delta(x) f(x) dx = f(0)$. Per dimostrare che $\delta(f)$ è illimitato è sufficiente considerare una successione quale

$$f_n(x) = \begin{cases} 0 & x > 1/n \\ 1 - nx & 0 < x < 1/n \\ f(-x) & x < 0 \end{cases}$$

Si trova $\delta(f) = 1$ e $\|f_n\|^2 = 2 \int_0^{1/n} (1 - nx)^2 dx \propto n^{-1}$, e dunque il rapporto $|\delta(f)|/\|f\|$ può assumere valori grandi a piacere. Un altro modo di verificare questo fatto è di considerare lo spazio $\ker(\delta)$ che risulta denso in $L_2(-1, 1)$ e quindi il suo complemento ortogonale risulta vuoto. ◀

Il risultato è del tutto generale.

Teor.4.1.5

TEOREMA 4.1.5. *Sia $\phi \in \mathfrak{H}^*$ un funzionale lineare illimitato. Allora il suo nucleo $\ker(\phi) = \{\xi \in D_\phi \mid \phi(\xi) = 0\}$ è denso in D_ϕ .*

PROVA. Se ϕ è illimitato esiste una successione $\eta_n \in D_\phi$ tale che

$$\phi(\eta_n) = n \|\eta_n\|$$

Allora dato un qualunque vettore $\xi \in D_\phi$ la successione

$$\xi_n = \xi - \frac{\phi(\xi)}{n \|\eta_n\|} \eta_n$$

tende a ξ per $n \rightarrow \infty$ e inoltre $\phi(\xi_n) = 0$. ■

✎ Ancora più generale è l'ovvio corollario del teorema precedente. In un qualunque intorno di un qualunque vettore $\xi \in D_\phi$ il funzionale non limitato ϕ assume qualunque valore complesso. Lasciamo la dimostrazione per esercizio. ◀

ESEMPIO 4-8. Come ultimo esempio si consideri il funzionale $\rho(\cdot)$ definito in $L_2(-\infty, \infty)$ da[‡]

$$\rho(f) = \int_{-\infty}^{\infty} f(x) dx, \quad \mathcal{D}(\rho) = L_2 \cap L_1$$

È facile costruire una successione di funzioni $\{f_n\}$ che mostrano come ρ sia illimitato. Più avanti incontreremo l'operatore

$$(\mathcal{F}f)(x) = \int_{-\infty}^{+\infty} e^{ixy} f(y) dy$$

(la *trasformata di Fourier*; è chiaro che $\rho(f) = \delta(\mathcal{F}f)$). ◀

4.2. Operatori lineari nello spazio di Hilbert

La semplicità della classificazione degli operatori lineari negli spazi a dimensione finita non si trasporta purtroppo al caso dello spazio di Hilbert. In particolare le proprietà spettrali degli operatori lineari sono estremamente più complesse - vedremo che esistono operatori con spettro vuoto oppure con spettro che riempie l'intero piano complesso - e di conseguenza il problema di definire una funzione di un operatore lineare sarà estremamente più impegnativo. Iniziamo lo studio degli operatori lineari nello spazio di Hilbert individuando alcune classi particolari di operatori.

4.2.1. Operatori lineari chiusi, continui, limitati. Un operatore lineare nello spazio di Hilbert è definito nel modo che segue:

Def.4.2.1

DEF. 4.2.1. Sia $D_{\mathcal{A}} \subset \mathfrak{H}$ un sottospazio lineare e $\mathcal{A} : D_{\mathcal{A}} \mapsto \mathfrak{H}$ una applicazione che soddisfi alle usuali condizioni di linearità, $\mathcal{A}(\alpha \xi + \beta \eta) = \alpha \mathcal{A}\xi + \beta \mathcal{A}\eta$. Allora la coppia $(\mathcal{A}, D_{\mathcal{A}})$ è un operatore lineare. $D_{\mathcal{A}}$ è detto il *dominio* dell'operatore. L'insieme dei vettori immagine sotto l'applicazione \mathcal{A}

$$\mathcal{R}_{\mathcal{A}} = \{\eta \in \mathfrak{H} \mid \eta = \mathcal{A}\xi, \xi \in D_{\mathcal{A}}\}$$

è detto il *codominio* (l'inglese "range") dell'operatore.

✎ Anche se comunemente ci si riferisce "all'operatore \mathcal{A} ", bisogna tenere ben presente che $D_{\mathcal{A}}$ costituisce una caratteristica importantissima dell'operatore, spesso determinante relativamente alle sue proprietà spettrali. ◀

[‡] $f \in L_1$ se f è assolutamente integrabile sulla retta reale; $f \in L_2 \cap L_1$ se esiste l'integrale $\int_{-\infty}^{\infty} |f(x)|^v dx$ sia per $v = 1$ che $v = 2$.

Considereremo in genere operatori con dominio denso in \mathfrak{H} - si dirà che l'operatore è *densamente definito*.

DEF. 4.2.2. Un operatore $(\mathcal{A}', D_{\mathcal{A}'})$ si dice un'estensione di $(\mathcal{A}, D_{\mathcal{A}})$ se

Def.4.2.2

$$(4.1) \quad \begin{aligned} D_{\mathcal{A}} &\subset D_{\mathcal{A}'} \\ \mathcal{A}'\xi &= \mathcal{A}\xi \quad \forall \xi \in D_{\mathcal{A}} \end{aligned}$$

e si indicherà per brevità $\mathcal{A}' \supset \mathcal{A}$. Un operatore ammette in generale diverse estensioni indipendenti a meno che $D_{\mathcal{A}}$ non coincida con \mathfrak{H} .

DEF. 4.2.3. L'insieme dei punti di $\mathfrak{H} \times \mathfrak{H}$

Def.4.2.3

$$\mathfrak{G}_{\mathcal{A}} = \{(\xi, \mathcal{A}\xi), \xi \in D_{\mathcal{A}}\}$$

è detto il *grafico* dell'operatore $(\mathcal{A}, D_{\mathcal{A}})$.

⚡ Si riserva la notazione $\langle x, y \rangle$ per indicare la "coppia" mentre useremo sempre $\langle x, y \rangle$ (ovvero la notazione di Dirac $\langle x|y \rangle$) per indicare il prodotto interno. ◀

⚡ Un purista potrebbe osservare che non c'è distinzione tra una funzione e il suo grafico (vedi ad es. [Hal82]): la funzione $f(x) = x^2$, in un certo senso, è l'insieme $\{(x, x^2) \in \mathbb{R}^2 \mid x \in \mathbb{R}\}$. Tuttavia la terminologia permette di evitare lunghi giri di parole. ◀

DEF. 4.2.4. Un operatore si dice *chiuso* se il suo grafico è un sottospazio chiuso di $\mathfrak{H} \times \mathfrak{H}$.

Def.4.2.4

Quest'ultima definizione è forse ingannevole nella sua apparente semplicità. Per esplicitarla, consideriamo una successione $\xi_1, \xi_2, \dots, \xi_n, \dots$ contenuta in $D_{\mathcal{A}}$ e la corrispondente successione

$$(\xi_n, \mathcal{A}\xi_n) \in \mathfrak{G}_{\mathcal{A}}, \quad n = 1, 2, 3, \dots$$

Se questa ultima converge, quindi se esistono i due limiti

$$\lim_{n \rightarrow \infty} \xi_n = \xi, \quad \lim_{n \rightarrow \infty} \mathcal{A}\xi_n = \eta,$$

allora se l'operatore è chiuso, ne segue che $(\xi, \eta) \in \mathfrak{G}$ e quindi $\xi \in D_{\mathcal{A}}$ e $\mathcal{A}\xi = \eta$. Ciò non esclude che esistano successioni ξ_n con limite $\xi \in D_{\mathcal{A}}$ ma $\mathcal{A}\xi_n$ divergente, oppure successioni con $\mathcal{A}\xi_n$ convergente ma $\lim \xi_n \notin D_{\mathcal{A}}$, oppure ancora successioni con $\mathcal{A}\xi_n$ convergente ma ξ_n divergente. La prima eventualità si presenta per operatori chiusi ma discontinui, l'ultima è sfruttata nella definizione di spettro continuo. Se $\mathfrak{G}_{\mathcal{A}}$ non è chiuso, si può esaminare la sua *chiusura* $\overline{\mathfrak{G}_{\mathcal{A}}}$: se questa costituisce un sottospazio lineare di $\mathfrak{H} \times \mathfrak{H}$ allora $\overline{\mathfrak{G}_{\mathcal{A}}}$ costituisce il grafico di un operatore $(\overline{\mathcal{A}}, D_{\overline{\mathcal{A}}})$ che viene detto la *chiusura* di \mathcal{A} .

DEF. 4.2.5. Si assuma che la chiusura del grafico $\mathfrak{G}_{\mathcal{A}}$ sia un sottospazio lineare di $\mathfrak{H} \times \mathfrak{H}$. Allora l'operatore avente grafico $\overline{\mathfrak{G}_{\mathcal{A}}}$ è detto la *chiusura* di $(\mathcal{A}, D_{\mathcal{A}})$

Def.4.2.5

DEF. 4.2.6. Un operatore $(\mathcal{A}, D_{\mathcal{A}})$ si dice *continuo* se per ogni successione $\xi_n \in D_{\mathcal{A}}$ avente limite in $D_{\mathcal{A}}$ vale la proprietà

Def.4.2.6

$$\mathcal{A} \lim_{n \rightarrow \infty} \xi_n = \lim_{n \rightarrow \infty} \mathcal{A}\xi_n$$

DEF. 4.2.7. Un operatore $(\mathcal{A}, D_{\mathcal{A}})$ si dice *limitato* se esiste un numero positivo M tale che

Def.4.2.7

$$\|\mathcal{A}\xi\| \leq M \|\xi\|.$$

Si chiama *norma* di \mathcal{A} il limite superiore (se esiste)

$$\|\mathcal{A}\| = \sup_{\xi \in D_{\mathcal{A}}} \frac{\|\mathcal{A}\xi\|}{\|\xi\|}.$$

(si veda def. 3.1.2).

Teor.4.2.1

TEOREMA 4.2.1. *Le due definizioni precedenti sono equivalenti; e cioè un operatore è limitato se e solo se è continuo.*

Omettiamo la dimostrazione che ricalca quella relativa ai funzionali lineari (Teorema 4.1.3).

Teor.4.2.2

TEOREMA 4.2.2 (Teorema del grafico chiuso). *Ogni operatore chiuso definito in tutto lo spazio di Hilbert è limitato (per la dimostrazione, non del tutto elementare, si veda [Tay58]).*

🔗 Quest'ultimo teorema spiega la necessità di definire operatori lineari in un opportuno dominio denso ma che non coincide con l'intero spazio. Vedremo che un operatore autoaggiunto è necessariamente chiuso, e d'altra parte si incontrano spesso nelle applicazioni operatori non-limitati. Quantunque sia possibile in linea di principio definire un operatore in tutto lo spazio sarebbe necessario rinunciare a qualche proprietà importante. Una motivazione più soddisfacente è che la definizione del dominio di un operatore contiene in molti casi informazioni sulle condizioni al contorno del problema fisico (tipico il caso di operatori differenziali in una regione limitata). ◀

4.2.1.1. *Criterio di Schur.* Un criterio utile per determinare se un operatore è limitato è formulato nel seguente

Teor.4.2.3

TEOREMA 4.2.3 (Schur). *Sia $[A_{mn}]_{m,n=1}^{\infty}$ una matrice infinita con elementi di matrice reali non-negativi. La matrice rappresenta un operatore limitato in ℓ_2 se esistono due successioni di numeri positivi $\{p_n | n = 1, 2, \dots\}$, $\{q_n | n = 1, 2, \dots\}$ e due numeri positivi α_1, α_2 tali che siano verificate le disuguaglianze*

$$(4.2) \quad \begin{aligned} \sum_{n=1}^{\infty} A_{mn} p_n &\leq \alpha_1 q_m \\ \sum_{m=1}^{\infty} A_{mn} q_m &\leq \alpha_2 p_n \end{aligned}$$

Inoltre si ha $\|A\| \leq \sqrt{\alpha_1 \alpha_2}$.

PROVA. Si tratta di maggioreare $\sum_m |\sum_n A_{mn} f_n|^2$; scriviamo

$$\sum_n A_{mn} f_n = \sum_n \left(\sqrt{A_{mn}} \sqrt{p_n} \right) \left(\sqrt{A_{mn}} f_n / \sqrt{p_n} \right)$$

Per la disuguaglianza di Schwartz segue

$$\left| \sum_n A_{mn} f_n \right|^2 \leq \sum_n (A_{mn} p_n) \sum_r (A_{mr} |f_r|^2 / p_r) \leq \alpha_1 q_m \sum_r A_{mr} |f_r|^2 / p_r.$$

Sommando su m (e invertendo l'ordine di sommazione) si ottiene il risultato cercato. ■

🔗 Dato che $|\sum_n A_{mn} f_n| \leq \sum_n |A_{mn}| |f_n|$, il criterio si applica a qualunque matrice a patto di prendere il valore assoluto degli elementi di matrice. Si vedano i Probl. 12-13-38-40-43. ◀

🔗 È interessante anche l'analogo continuo del criterio, applicabile a operatori del tipo

$$(Af)(x) = \int_a^b A(x, y) f(y) dy, \quad f \in L_2(a, b)$$

È sufficiente modificare la formula con le sostituzioni $\sum \rightarrow \int$, $n \rightarrow x$. ◀

4.2.2. Aggiunto Hermitiano di operatori limitati. Come nel caso finito-dimensionale, consideriamo il funzionale lineare

$$\phi_\eta(\xi) = \langle \eta | \mathcal{A} | \xi \rangle$$

Se \mathcal{A} è un operatore limitato, si avrà

$$|\phi_\eta(\xi)| = |\langle \eta | \mathcal{A} | \xi \rangle| \leq \|\eta\| \|\mathcal{A} \xi\| \leq M \|\eta\| \|\xi\|$$

dunque ϕ_η è un funzionale limitato ed esiste di conseguenza, per il teorema di Riesz, un unico vettore $\phi_\eta^\#$ tale che sia

$$\langle \eta | \mathcal{A} | \xi \rangle = \langle \phi_\eta^\# | \xi \rangle$$

La corrispondenza che mappa $\eta \mapsto \phi_\eta^\#$ si vede facilmente essere lineare e dunque si definisce

$$\phi_\eta^\# = \mathcal{A}^\dagger \eta$$

\mathcal{A}^\dagger è un operatore limitato, con la stessa norma di \mathcal{A} , che viene chiamato *l'aggiunto Hermitiano* di \mathcal{A} . Se $\mathcal{A} \equiv \mathcal{A}^\dagger$ allora si dice che, come nel caso finito-dimensionale, che \mathcal{A} è *autoaggiunto*.

4.2.3. Operatori isometrici e unitari.

DEF. 4.2.8. Un operatore $(\mathcal{U}, D_{\mathcal{U}})$ è detto *isometrico* se

Def.4.2.8

$$\langle \mathcal{U} \xi, \mathcal{U} \eta \rangle = \langle \xi, \eta \rangle$$

per ogni coppia di vettori ξ, η in $D_{\mathcal{U}}$.

Ogni operatore isometrico è ovviamente limitato con norma uno. Se \mathcal{U} è isometrico e densamente definito esso può essere esteso per continuità a tutto lo spazio di Hilbert. Dalla definizione segue immediatamente

$$\mathcal{U}^\dagger \mathcal{U} = \mathbb{1}$$

di conseguenza, da $\mathcal{U} \xi = \mathcal{U} \eta$ segue $\xi = \eta$. Dunque un operatore isometrico definisce una applicazione biunivoca dello spazio di Hilbert sul *codominio* (o *range*) $\mathcal{R}_{\mathcal{U}}$ [§]. Si distingue il caso in cui $\mathcal{R}_{\mathcal{U}}$ è denso in \mathfrak{H} dal caso in cui $\mathcal{R}_{\mathcal{U}}$ non lo è. In quest'ultimo caso, esiste almeno un vettore η ortogonale a tutti i vettori $\mathcal{U} \xi$ e perciò $\mathcal{U}^\dagger \eta = 0$. In questa situazione si ha che $\mathcal{U}^\dagger \mathcal{U} \neq \mathcal{U} \mathcal{U}^\dagger$.

n \neq ∞ Il caso di operatori isometrici con range non denso in \mathfrak{H} è specifico dello spazio di Hilbert. In spazi a un numero finito di dimensioni infatti se \mathcal{U}^\dagger ammette un autovettore con autovalore zero la sua matrice rappresentativa è singolare e dunque anche \mathcal{U} avrebbe un autovalore zero (vedi il Teorema 3.3.5). ◀

DEF. 4.2.9. Un operatore isometrico \mathcal{U} si dice *unitario* se il suo codominio $\mathcal{R}_{\mathcal{U}}$ è denso in \mathfrak{H} .

Def.4.2.9

Un operatore unitario ammette dunque un inverso densamente definito e inoltre $\mathcal{U}^{-1} = \mathcal{U}^\dagger$, per cui $\mathcal{U} \mathcal{U}^\dagger = \mathcal{U}^\dagger \mathcal{U} = \mathbb{1}$.

ESEMPIO 4-9. Sia \mathcal{U} l'operatore definito in tutto ℓ_2 da

$$\mathcal{U} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ z_1 \\ z_2 \\ \vdots \\ z_{n-1} \\ \vdots \end{pmatrix}$$

[§]Ricordiamo che il codominio di un operatore \mathcal{U} è dato dall'insieme $\{\eta \in \mathfrak{H} \mid \exists \xi \ni \eta = \mathcal{U} \xi\}$.

Dato che le componenti del vettore vengono semplicemente avanzate di un posto è ovvio che la norma del vettore rimane invariata, dunque \mathcal{U} è isometrico. Tuttavia il codominio di \mathcal{U} è dato dai vettori ortogonali a

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix}$$

e perciò \mathcal{U} non è unitario. In effetti il suo aggiunto Hermitiano è dato da

$$\mathcal{U}^\dagger \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_n \\ \vdots \end{pmatrix} = \begin{pmatrix} z_2 \\ z_3 \\ z_4 \\ \vdots \\ z_{n+1} \\ \vdots \end{pmatrix}$$

Si verifica perciò che $\mathcal{U}\mathcal{U}^\dagger = \mathbb{1} - |1\rangle\langle 1|$. ◀

🔗 *Operatori isometrici non-unitari si presentano in teoria della diffusione in Meccanica Quantistica (Operatori di Möller, si veda [New66]).* ◀

ESEMPIO 4-10. Sia $f(x) \in L_2(-\infty, \infty)$; per ogni numero reale a sia definito l'operatore

$$(\mathcal{T}_a f)(x) \equiv f(x - a)$$

che rappresenta la traslazione della funzione f . È immediato verificare che \mathcal{T}_a è isometrico. Inoltre, dato che $\mathcal{T}_a \mathcal{T}_{-a} = \mathbb{1}$, \mathcal{T}_a risulta invertibile e pertanto unitario. ◀

ESEMPIO 4-11. Sia $f(x) \in L_2(0, \infty)$; per ogni numero reale a sia definito l'operatore

$$(\mathcal{T}_a f)(x) \equiv \begin{cases} f(x - a) & (x > a) \\ 0 & (0 < x < a) \end{cases}$$

che rappresenta ancora una traslazione della funzione f ; dato che siamo sulla semiretta $x > 0$ è come considerare la funzione identicamente nulla per x negativi. È immediato verificare che \mathcal{T}_a è isometrico se $a > 0$. Tuttavia, dato che per a negativo la traslazione avviene verso sinistra, la funzione viene “tagliata” della porzione $0 < x < -a$ e di conseguenza l'operatore \mathcal{T}_a per $a < 0$ non è isometrico e ammette un intero sottospazio ∞ -dimensionale di vettori con $\mathcal{T}_a \xi = 0$. ◀

ESEMPIO 4-12. Più un generale si consideri in $L_2(\mathbb{R}^3)$ l'insieme di operatori definiti da

$$(\mathcal{T}_{R,a} f)(x) = f(R^{-1}(x - a))$$

dove R è una qualunque matrice ortogonale, che rappresenta una rotazione nello spazio \mathbb{R}^3 . Anche in questo caso si verifica facilmente che \mathcal{T} è isometrico e invertibile, dunque unitario. Ciò discende dal fatto che lo Jacobiano della trasformazione $y = Rx + a$ è dato da $\det R = \pm 1$. ◀

4.2.4. Aggiunto Hermitiano per operatori non limitati. Sia $(\mathcal{A}, D_{\mathcal{A}})$ un operatore non limitato definito densamente in \mathfrak{H} . Consideriamo il funzionale lineare

$$\Phi_\eta(\xi) = \langle \eta, \mathcal{A}\xi \rangle, \quad \xi \in D_{\mathcal{A}}.$$

Sia $D_{\mathcal{A}^\dagger}$ il sottospazio di \mathfrak{H} tale che per ogni $\eta \in D_{\mathcal{A}^\dagger}$ esiste un vettore $\Phi_\eta^\#$ che soddisfa l'identità

$$\Phi_\eta(\xi) = \langle \eta, \mathcal{A}\xi \rangle = \langle \Phi_\eta^\#, \xi \rangle.$$

Si verifica facilmente che la corrispondenza $\eta \mapsto \Phi_\eta^\sharp$ è lineare. Si definisce allora l'operatore \mathcal{A}^\dagger , l'aggiunto Hermitiano di \mathcal{A} come l'operatore lineare con dominio $D_{\mathcal{A}^\dagger}$ tale che $\mathcal{A}^\dagger \eta = \Phi_\eta^\sharp$. Si noti che a priori non si può dire nulla in generale sul dominio $D_{\mathcal{A}^\dagger}$ in quanto manca la possibilità di invocare il teorema di Riesz. In effetti si danno esempi di operatori non limitati il cui aggiunto ha dominio ristretto al vettore nullo $\mathbf{0}$.

Perchè la definizione sia consistente si tratta di mostrare che la corrispondenza è effettivamente lineare. Si procede come segue: se $\langle \eta, \mathcal{A} \xi \rangle = \langle \Phi_\eta, \xi \rangle = \langle \Phi_\eta, \xi \rangle$ ne segue $\langle \Phi_\eta - \Phi_\eta, \xi \rangle = 0$ e dunque $\Phi_\eta = \Phi_\eta$ in quanto $D_{\mathcal{A}}$ è denso in \mathfrak{H} .

Ricordando la Def. 4.2.2, si ha

TEOREMA 4.2.4. *Sia $\mathcal{A} \supset \mathcal{B}$ allora $\mathcal{B}^\dagger \supset \mathcal{A}^\dagger$.*

Teor.4.2.4

PROVA. Le condizioni che definiscono $D_{\mathcal{B}^\dagger}$ sono senz'altro soddisfatte in $D_{\mathcal{A}^\dagger}$. ■

È interessante considerare il grafico (secondo la Def. 4.2.3) dell'operatore aggiunto. A questo scopo, definiamo l'operatore \mathcal{J} in $\mathfrak{H} \times \mathfrak{H}$ che agisce come segue

$$\mathcal{J}(\xi, \eta) = (i\eta, -i\xi).$$

\mathcal{J} è un operatore unitario e $\mathcal{J}^2 = \mathbb{1}$. Applichiamo \mathcal{J} a tutti i punti del grafico $G_{\mathcal{A}}$ di \mathcal{A} e consideriamo il complemento ortogonale dell'immagine di $G_{\mathcal{A}}$ sotto l'azione di \mathcal{J} ; si ha

TEOREMA 4.2.5. *Il grafico di $(\mathcal{A}^\dagger, D_{\mathcal{A}^\dagger})$ coincide con $(\mathcal{J} G_{\mathcal{A}})^\perp$.*

Teor.4.2.5

PROVA. $(\mathcal{J} G_{\mathcal{A}})^\perp$ consiste di tutte le coppie (η, Φ) tali che

$$\langle \mathcal{J}(\xi, \mathcal{A} \xi), (\eta, \Phi) \rangle = 0$$

ma ciò equivale a

$$\langle (i\mathcal{A}\xi, -i\xi), (\eta, \Phi) \rangle = \langle i\mathcal{A}\xi, \eta \rangle + \langle -i\xi, \Phi \rangle = 0$$

che coincide con la definizione di aggiunto Hermitiano. ■

Ne discende come corollario immediato

TEOREMA 4.2.6. *$(\mathcal{A}^\dagger, D_{\mathcal{A}^\dagger})$ è un operatore chiuso.*

Teor.4.2.6

TEOREMA 4.2.7. *L'aggiunto Hermitiano di un operatore chiuso è densamente definito.*

Teor.4.2.7

PROVA. Assumiamo che $D_{\mathcal{A}^\dagger}$ non sia denso. Allora se $\eta \in D_{\mathcal{A}^\dagger}^\perp$ segue

$$\langle (\eta, 0), (\xi, \mathcal{A}^\dagger \xi) \rangle = 0, \quad \forall \xi \in D_{\mathcal{A}^\dagger}$$

dunque $(\eta, 0)$ è ortogonale a $G_{\mathcal{A}^\dagger}$ ed appartiene di conseguenza alla chiusura di $\mathcal{J} G_{\mathcal{A}}$. Da $(0, \eta) \in G_{\mathcal{A}}$ segue l'esistenza di un vettore ξ tale che $(0, \eta) = (\xi, \mathcal{A} \xi)$ che implica $\eta = 0$. ■

DEF. 4.2.10. Un operatore \mathcal{A} è detto *simmetrico* se

Def.4.2.10

$$\langle \xi, \mathcal{A} \eta \rangle = \langle \mathcal{A} \xi, \eta \rangle$$

per ogni coppia di vettori in $D_{\mathcal{A}}$.

TEOREMA 4.2.8. *Per un operatore simmetrico $(\mathcal{A}, D_{\mathcal{A}})$ vale la relazione*

Teor.4.2.8

$$\boxed{\mathcal{A}^\dagger \supseteq \mathcal{A}}$$

cioè l'aggiunto Hermitiano di \mathcal{A} è una sua estensione, oppure coincide con \mathcal{A} .

TEOREMA 4.2.9. *L'operatore simmetrico \mathcal{B} sia un'estensione dell'operatore simmetrico \mathcal{A} ; allora*

Teor.4.2.9

$$\boxed{\mathcal{A} \subset \mathcal{B} \subseteq \mathcal{B}^\dagger \subseteq \mathcal{A}^\dagger}$$

Def.4.2.11

DEF. 4.2.11. Un operatore $(\mathcal{A}, D_{\mathcal{A}})$ è detto *autoaggiunto* se coincide con il suo aggiunto Hermitiano, e cioè

$$\mathcal{A} = \mathcal{A}^\dagger, \quad D_{\mathcal{A}} = D_{\mathcal{A}^\dagger}$$

Def.4.2.12

DEF. 4.2.12. Un'operatore simmetrico è detto *essenzialmente autoaggiunto* se $(\mathcal{A}^\dagger)^\dagger = \mathcal{A}^\dagger$ ossia se il suo aggiunto Hermitiano è autoaggiunto, nel qual caso \mathcal{A}^\dagger è l'unica estensione autoaggiunta di \mathcal{A} .

4.2.5. Operatori differenziali in $L_2(a, b)$. Sia $f \in L_2(a, b)$ continua, differenziabile con derivata in L_2 . Allora

$$(\mathcal{D}f)(x) = -i \frac{df(x)}{dx}$$

definisce un operatore lineare. L'equazione precedente specifica l'azione dell'operatore \mathcal{D} mentre le specifiche sulla funzione f ne definiscono il dominio. La stessa espressione differenziale corrisponde a tanti differenti operatori lineari a seconda di come sia specificato il dominio (ad es. potremmo richiedere che f sia C^∞ , oppure analitica, oppure che si annulli al bordo, etc.). È necessario familiarizzarsi con il fatto che la scelta del dominio è spesso decisiva per caratterizzare le proprietà dell'operatore. Ciò corrisponde d'altronde, nelle applicazioni alla Fisica, al fatto che le equazioni differenziali (eq. d'onda, eq. di Schroedinger, eq. del calore, etc.) ammettono in generale una varietà di soluzioni ma sono le *condizioni al contorno* del problema che, ove siano "ben poste", permettono di individuare la soluzione corretta in modo univoco. L'insistenza sull'importanza della definizione di dominio di un operatore differenziale non è dunque un puro scrupolo di rigore matematico (per quanto legittimo) ma è dettata, al contrario, dalle esigenze delle applicazioni. Studieremo ora l'esempio più semplice di operatore differenziale, rimandando agli esercizi per altri esempi.

4.2.5.1. *Operatori differenziali del I ordine.* Sia \mathfrak{C}^∞ il dominio di $L_2(a, b)$ così definito[¶]

$$\mathfrak{C}^\infty = \{f \in L_2(a, b) \mid f \text{ derivabile infinite volte, } f^{(n)}(a) = f^{(n)}(b) = 0 \ (\forall n \geq 0)\}$$

Sia $(\mathcal{D}_0, \mathfrak{C}^\infty)$ l'operatore definito nel dominio \mathfrak{C}^∞ da

$$(\mathcal{D}_0 f)(x) = -i \frac{df(x)}{dx}.$$

Teor.4.2.10

TEOREMA 4.2.10. $(\mathcal{D}_0, \mathfrak{C}^\infty)$ è un operatore simmetrico ma non autoaggiunto.

PROVA. Per ogni $f \in \mathfrak{C}^\infty$ si ha

$$(4.3) \quad \langle f, \mathcal{D}_0 g \rangle = -i \int_a^b \overline{f(x)} g'(x) dx = i \int_a^b \overline{f'(x)} g(x) dx = \langle \mathcal{D}_0 f, g \rangle$$

(niente termini al contorno nell'integrazione per parti per la definizione di \mathfrak{C}^∞). L'operatore aggiunto \mathcal{D}_0^\dagger si costruisce come segue: si cercano i vettori $f \in L_2$ tali che esista un $h \in L_2$ che soddisfi la relazione

$$\langle f, \mathcal{D}_0 g \rangle = \langle h, g \rangle \ (\forall g \in \mathfrak{C}^\infty)$$

Esplicitamente

$$(4.4) \quad -i \int_a^b \overline{f(x)} g'(x) dx = \int_a^b \overline{h(x)} g(x) dx$$

[¶]Qui stiamo considerando funzioni continue con tutte le derivate, a cui si applica il teorema fondamentale del calcolo integrale, che non è applicabile in generale ai vettori di L_2 .

Ora, ogni $g \in \mathfrak{C}^\infty$ può rappresentarsi nella forma $g(x) = \int_a^x z(y) dy$ che inserita nell'Eq.4.4 fornisce

$$\begin{aligned} -i \int_a^b \overline{f(x)} z(x) dx &= \int_a^b \overline{h(x)} \int_a^x z(y) dx dy \\ &= \int_a^b z(y) \int_y^b \overline{h(x)} dx dy \end{aligned}$$

ossia

$$(4.5) \quad \int_a^b z(x) \left[-i \overline{f(x)} - \int_x^b \overline{h(y)} dy \right] dx = 0.$$

Ora, $z(x)$ non individua un sottospazio denso, infatti se il suo integrale è in \mathfrak{C}^∞ , essa deve soddisfare il vincolo

$$\int_a^b z(x) dx = g(b) - g(a) = 0$$

e dunque $z(x)$ è ortogonale alla funzione costante. Dall'Eq. (4.5) si conclude allora che

$$f(x) = -i \int_x^b h(x) dy + f(b)$$

ed infine

$$h(x) = -i \frac{df(x)}{dx}.$$

Si conclude che il dominio di \mathcal{D}_0^\dagger è costituito da tutte le *funzioni assolutamente continue* ossia esprimibili come l'integrale indefinito di una qualunque $h \in L_2$ avente derivata in L_2 . Il dominio dell'operatore aggiunto non prevede perciò le condizioni al contorno di annullamento della funzione con tutte le sue derivate che caratterizzano il dominio di \mathcal{D}_0 . Abbiamo verificato che $\mathcal{D}_{\mathcal{D}_0} \subset \mathcal{D}_{\mathcal{D}_0^\dagger}$. ■

Definiamo ora una estensione di \mathcal{D}_0 definita da ^{||}

$$\begin{aligned} (\mathcal{D}_* f)(x) &= -i \frac{df(x)}{dx} \\ f &\in \mathfrak{C}_* = \{f \in L_2(a, b), f(x) \text{ a.c.}, df/dx \in L_2, f(b) = f(a) = 0\} \end{aligned}$$

\mathcal{D}_* è una estensione simmetrica di \mathcal{D}_0 ma non ancora un operatore autoaggiunto. Ulteriori estensioni sono definite alleggerendo ancora le condizioni al contorno, ma preservando la proprietà di simmetria. Sia (per $|\alpha| \leq \pi$)

$$\begin{aligned} (\mathcal{D}_\alpha f)(x) &= -i \frac{df(x)}{dx} \\ f &\in \mathfrak{C}_\alpha = \{f \in L_2(a, b), f(x) \text{ a.c.}, df/dx \in L_2, f(b) = f(a) e^{i\alpha}\} \end{aligned}$$

Dato che $\mathfrak{C}_\alpha \supset \mathfrak{C}^\infty$ si ha $\mathcal{D}_\alpha^\dagger \subset \mathcal{D}_0^\dagger$ ed è perciò superfluo ripetere tutto l'argomento di Pag. 120. Sarà sufficiente verificare che le condizioni al contorno su $\mathcal{D}_\alpha^\dagger$ coincidono con quelle su \mathcal{D}_α per concludere che \mathcal{D}_α è autoaggiunto per ogni valore reale prefissato per α . Si ha infatti

$$\langle g, \mathcal{D}_\alpha f \rangle = -i \overline{g(x)} f(x) \Big|_a^b + i \int_a^b \overline{g'(x)} f(x) dx = \langle -ig', f \rangle + i(\overline{g(b)} e^{i\alpha} - \overline{g(a)}) f(a)$$

Ciò implica $g(b) = g(a) e^{i\alpha}$ e dunque le stesse condizioni al contorno su g che su f .

^{||}Conveniamo che "a.c." significa d'ora in poi "assolutamente continua". Le due condizioni $f \in L_2$ e f a.c. sono indipendenti.

In conclusione: $(\mathcal{D}_0, \mathcal{C}^\infty)$ è un operatore simmetrico che ammette un numero infinito di estensioni autoaggiunte \mathcal{D}_α .

Una tecnica sistematica che permette di decidere se un operatore simmetrico ammetta estensioni autoaggiunte e che fornisce anche un procedimento per costruirle esplicitamente sarà esposta succintamente più avanti (“spazi di difetto”, vedi Sez.4.3.6).

Un operatore differenziale è in generale definito da una espressione del tipo

$$\mathcal{P}(x, d/dx) = \sum_{j=0}^n a_j(x) \left(\frac{d}{dx} \right)^j$$

e dalla scelta di un dominio $\mathcal{D}_\mathcal{P} \subset L_2(a, b)$ che sarà in genere costituito da funzioni derivabili $n-1$ volte con derivata $(n-1)$ -esima a.c., $\in L_2$ e con condizioni aggiuntive che vincolino tra loro attraverso equazioni lineari i valori che la funzione e le sue derivate assumono al bordo. La teoria generale di tali operatori è trattata ampiamente in [Nai68]. Ci limiteremo nel seguito a considerare un esempio molto importante.

4.2.5.2. *L'operatore $(d/dx)^2$ in $L_2(a, b)$.* Sia \mathcal{H}_0 definito dall'espressione differenziale

$$(\mathcal{H}_0 f)(x) = -\frac{d^2 f(x)}{dx^2}$$

con dominio

$$\mathcal{D}_{\mathcal{H}_0} = \{f \text{ a.c.}, f' \text{ a.c.}, f'' \in L_2, f(a) = f(b) = 0, f'(a) = f'(b) = 0\}$$

Teor.4.2.11

TEOREMA 4.2.11. $(\mathcal{H}_0, \mathcal{D}_{\mathcal{H}_0})$ è simmetrico, non autoaggiunto e ammette infinite estensioni autoaggiunte.

PROVA. La simmetria segue da una doppia applicazione dell'integrazione per parti. Per costruire l'aggiunto Hermitiano di \mathcal{H}_0 procediamo così: sia $z(x) \in L_2$ con

$$\int_a^b z(x) (A + Bx) dx = 0 \quad (\forall A, B \in \mathbb{C}).$$

Allora $f(x) = \int_a^x (x-y) z(y) dy$ è la più generale espressione di una funzione in $\mathcal{D}_{\mathcal{H}_0}$. Analogamente a quanto fatto in precedenza nel § 4.2.5.1 si trova che l'equazione

$$\langle g, \mathcal{H}_0 f \rangle = \langle h, f \rangle$$

implica

$$\int_a^b \overline{z(x)} \left[g(x) - \int_x^b (y-x) h(y) dy \right] dx = 0$$

e quindi $\mathcal{D}_{\mathcal{H}_0^\dagger}$ è costituito dalle funzioni del tipo

$$g(x) = \int_x^b (y-x) h(y) dy + Ax + B$$

e cioè $\mathcal{D}_{\mathcal{H}_0^\dagger}$ si ottiene da $\mathcal{D}_{\mathcal{H}_0}$ eliminando le condizioni al contorno. Per costruire estensioni autoaggiunte di \mathcal{H}_0 dovremo rilasciare alcune delle quattro condizioni lineari al bordo. Poniamo per esempio (“c.c.” = *condizioni al contorno*)

$$(\mathcal{H} f)(x) = -f''(x)$$

$$\mathcal{D}_\mathcal{H} = \{f \text{ a.c.}, f' \text{ a.c.}, f'' \in L_2, \& \text{c.c.}\}$$

dove

$$\text{c.c.} = \begin{cases} \alpha_1 f(a) + \alpha_2 f(b) + \alpha_3 f'(a) + \alpha_4 f'(b) = 0 \\ \beta_1 f(a) + \beta_2 f(b) + \beta_3 f'(a) + \beta_4 f'(b) = 0 \end{cases}$$

Dato che $D_{\mathcal{H}} \supset D_{\mathcal{H}_0}$ si avrà $D_{\mathcal{H}^\dagger} \subset D_{\mathcal{H}_0}$ e quindi potremo limitarci ad imporre che le condizioni al contorno per \mathcal{H}^\dagger siano identiche a quelle di \mathcal{H} :

$$\langle g, 'Hf \rangle = \langle -g'', f \rangle + [\overline{g(x)} f'(x) - \overline{g'(x)} f(x)]|_a^b$$

e sappiamo già che $\mathcal{H}^\dagger g = H_0^\dagger g = -g''$. Si tratta di determinare la matrice

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{pmatrix}$$

in modo che le due condizioni

$$(4.6) \quad \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} f(a) \\ f(b) \\ f'(a) \\ f'(b) \end{pmatrix} = 0, \quad \overline{g(b)} f'(b) - \overline{g'(b)} f(b) = \overline{g(a)} f'(a) - \overline{g'(a)} f(a)$$

implichino la stessa condizione sulla $g(x)$, cioè

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{pmatrix} \begin{pmatrix} g(a) \\ g(b) \\ g'(a) \\ g'(b) \end{pmatrix} = 0$$

La matrice $\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$ deve avere rango due in modo che le condizioni al contorno siano due equazioni linearmente indipendenti. Inoltre due matrici $\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$ e $\begin{pmatrix} \alpha'_i \\ \beta'_i \end{pmatrix}$ che siano legate da una relazione del tipo

$$\begin{pmatrix} \alpha'_i \\ \beta'_i \end{pmatrix} = \begin{pmatrix} u & v \\ w & z \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$$

definiscono le stesse c.c. a patto che il determinante $uz - vw$ non si annulli. Possiamo sfruttare questo fatto per ridurre una parte della matrice a un formato standard. Ad es. se il minore

$$\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix}$$

ha determinante diverso da zero, si possono ridurre le c.c. alla forma

$$(4.7) \quad \begin{aligned} f(a) + \alpha_3 f'(a) + \alpha_4 f'(b) &= 0 \\ f(b) + \beta_3 f'(a) + \beta_4 f'(b) &= 0 \end{aligned}$$

che implicano, combinate con l'Eq. (4.6)

$$\begin{aligned} g(a) + \overline{\alpha_3} g'(a) - \overline{\beta_3} g'(b) &= 0 \\ g(b) - \overline{\alpha_4} g'(a) + \overline{\beta_4} g'(b) &= 0 \end{aligned}$$

Un'estensione autoaggiunta di \mathcal{H}_0 è perciò caratterizzata da condizioni al contorno date dall'Eq. (4.7) con α_3 e β_4 reali e $\alpha_4 = -\overline{\beta_3}$ (cioè quattro parametri reali). Condizioni simili si possono considerare nel caso che altri minori 2×2 della matrice generale nell'Eq. (4.6) risulti non singolare**. ■

Come casi particolari di c.c. autoaggiunte citiamo i seguenti:

- (a) $f(a) = f(b) = 0$ - (condizioni di annullamento o di Dirichlet)
- (b) $f'(a) = f'(b) = 0$ - (condizioni di Neumann)

**Il problema può formularsi in termini di "sottospazi lagrangiani" - si veda oltre nella sezione 4.3.6.

- (c) $f(b) = \exp(i\tau) f(a)$, $f'(b) = \exp(i\tau) f'(a)$ - (condizioni periodiche con sfasamento)
- (d) $f(a) = f'(b) = 0$ - (canna d'organo)
- (e) $f'(a) = k f(a)$, $f'(b) = h f(b)$.

Abbiamo qui considerato operatori differenziali definiti su di un intervallo finito (a, b) . Il caso in cui uno o entrambi gli estremi sono infiniti è ovviamente interessante per le applicazioni, così come è interessante lo studio di operatori differenziali su varietà differenziali, quali la sfera o il toro. Alcuni esempi saranno incontrati più oltre, ma una trattazione sistematica esula completamente dallo scopo di queste lezioni.

4.3. Teoria spettrale

La teoria spettrale degli operatori lineari nello spazio di Hilbert è fondata, come per gli spazi finito-dimensionali, sulla definizione di *risolvente* di un operatore. Sia $(\mathcal{A}, D_{\mathcal{A}})$ un operatore densamente definito nello spazio di Hilbert \mathfrak{H} e sia z un numero complesso tale che

$$*) \quad z\mathbb{1} - \mathcal{A} : D_{\mathcal{A}} \xrightarrow{1:1} R_{z\mathbb{1} - \mathcal{A}} \subseteq \mathcal{H}$$

ossia risulti una trasformazione biunivoca di $D_{\mathcal{A}}$ in $R_{z\mathbb{1} - \mathcal{A}}$ (cioè $(z\mathbb{1} - \mathcal{A})|\xi\rangle = 0$ solo se $|\xi\rangle = 0$). È definita allora la trasformazione inversa

$$(z\mathbb{1} - \mathcal{A})^{-1} : R_{z\mathbb{1} - \mathcal{A}} \mapsto D_{\mathcal{A}}$$

L'operatore $\mathcal{R}(z, \mathcal{A}) = (z\mathbb{1} - \mathcal{A})^{-1}$ è detto *il risolvente* di \mathcal{A} . L'insieme dei valori z per cui sono soddisfatte le due ulteriori condizioni

- ***) $R_{z\mathbb{1} - \mathcal{A}}$ è denso in \mathfrak{H} ;
- ****) $(z\mathbb{1} - \mathcal{A})^{-1}$ è limitato;

si chiama *l'insieme risolvente* e verrà indicato con $\text{RES}(\mathcal{A}) \subset \mathbb{C}$. Il set complementare $\mathbb{C} \setminus \text{RES}(\mathcal{A}) = \mathbb{C} \setminus \text{RES}(\mathcal{A})$ è chiamato *lo spettro* di \mathcal{A} e verrà indicato con $\sigma(\mathcal{A})$. Risulta utile individuare nello spettro varie componenti:

$\mathbf{P}\sigma(\mathcal{A})$: lo spettro *puntuale* (o *spettro discreto*)

$\mathbf{C}\sigma(\mathcal{A})$: lo spettro *continuo*

$\mathbf{p}\sigma(\mathcal{A})$: lo spettro *residuo*.

Lo *spettro discreto* è l'insieme dei valori complessi z per i quali cade la condizione (*) ossia per i quali esiste almeno un vettore non nullo nel dominio di \mathcal{A} con

$$\mathcal{A}|z\rangle = z|z\rangle$$

z si chiama in questo caso un autovalore di \mathcal{A} e $|z\rangle$ un autovettore appartenente all'autovalore z .

Lo *spettro continuo* di \mathcal{A} è l'insieme degli z tali che $z\mathbb{1} - \mathcal{A}$ soddisfa (*, **) ma non (***); in modo più esplicito, $z \in \mathbf{C}\sigma$ se esiste una successione di vettori $\{\xi_n, n = 1, 2, \dots, \infty\}$, con $\|\xi_n\| = 1$ per cui esista il limite

$$\lim_{n \rightarrow \infty} \|(z\mathbb{1} - \mathcal{A})\xi_n\| = 0$$

In questa situazione $(z\mathbb{1} - \mathcal{A})^{-1}$ esiste ma è illimitato.

Lo *spettro residuo* $\mathbf{p}\sigma(\mathcal{A})$ è l'insieme dei numeri complessi z per i quali è violata la condizione (**), cioè esiste almeno un vettore ortogonale al codominio $R_{z\mathbb{1} - \mathcal{A}}$. Questo significa che

$$\langle \eta, (z\mathbb{1} - \mathcal{A})\xi \rangle = 0 = \langle (\bar{z}\mathbb{1} - \mathcal{A}^\dagger)\eta, \xi \rangle$$

e dunque \bar{z} è nello spettro puntuale di \mathcal{A}^\dagger .

$n \neq \infty$ Notiamo che spettro continuo e spettro residuo sono caratteristici dello spazio di Hilbert (infinita dimensioni). Inoltre non è più semplicemente l'equazione agli autovalori $\mathcal{A}\xi = \lambda\xi$ a determinare lo spettro di un operatore, ma è necessaria un'analisi più raffinata. Infine, a differenza del caso finito-dimensionale, non si ha alcuna garanzia a priori che lo spettro non

sia vuoto. Vedremo infatti esempi di operatori aventi tutto il piano complesso come insieme risolvente, e che ammettono un'estensione avente spettro puntuale che invade tutto il piano complesso! È solo limitandosi a classi particolari di operatori (autoaggiunti, unitari, normali, compatti, etc.) che si potrà mettere un po' d'ordine nella teoria spettrale.

4.3.1. Prime proprietà ed esempi.

TEOREMA 4.3.1. *Sia \mathcal{A} un operatore limitato con norma M . Allora lo spettro di \mathcal{A} è contenuto nel cerchio $|z| \leq M$.*

Teor.4.3.1

PROVA. Sia $|z| > M$. Allora $(z\mathbb{1} - \mathcal{A})\xi = \mathbf{0}$ implica $\xi = \mathbf{0}$. Inoltre $R_{z\mathbb{1} - \mathcal{A}}$ è denso in \mathfrak{H} , in quanto anche \mathcal{A}^\dagger è limitato con la stessa norma M . Infine non può esistere alcuna successione che definisca z come punto di $\mathbf{C}\sigma(\mathcal{A})$ in quanto

$$\|(z\mathbb{1} - \mathcal{A})\xi_n\| \geq |z| \|\xi_n\| - \|\mathcal{A}\xi_n\| \geq (|z| - M)\|\xi_n\|$$

■

ESEMPIO 4-13. Sia \mathcal{T} l'operatore in ℓ_2 definito da

$$\mathcal{T}(z_1, z_2, \dots, z_n, \dots) = (z_2, z_3, \dots, z_n, \dots)$$

\mathcal{T} è chiamato l'operatore *di shift*. Si ha ovviamente $\|\mathcal{T}\| = 1$ e quindi $\text{RES}(\mathcal{T}) \supseteq \{|z| > 1\}$. L'equazione agli autovalori

$$\mathcal{T}(z_1, z_2, \dots, z_n, \dots) = (z_2, z_3, \dots, z_n, \dots) = \lambda(z_1, z_2, \dots, z_n, \dots)$$

ha per soluzione

$$|\lambda\rangle = \frac{1}{\sqrt{1 - |\lambda|^2}} (1, \lambda, \lambda^2, \lambda^3, \dots, \lambda^{n-1}, \dots)$$

che è in ℓ_2 per ogni valore complesso λ con $|\lambda| < 1$. Scegliendo una successione di vettori normalizzati $|\lambda_n\rangle$ con $\lambda_n \rightarrow e^{i\phi}$ si trova che i punti sulla circonferenza di raggio uno appartengono allo spettro continuo $\mathbf{C}\sigma(\mathcal{T})$. In conclusione $\mathbf{P}\sigma(\mathcal{T}) = \{|z| < 1\}$, $\mathbf{C}\sigma(\mathcal{T}) = \{|z| = 1\}$, $\rho(\mathcal{T}) = \emptyset$, $\text{RES}(\mathcal{T}) = \{|z| > 1\}$. ▶

ESEMPIO 4-14. Sia q l'operatore definito in $L_2(a, b)$ dalla espressione

$$(qf)(x) = xf(x)$$

Se entrambi gli estremi dell'intervallo sono finiti, allora q è limitato:

$$\int_a^b x^2 |f(x)|^2 dx \leq \max(a^2, b^2) \|f\|^2$$

L'operatore risolvente è dato semplicemente da

$$(\mathcal{R}(z, q)f)(x) = \frac{f(x)}{z - x}$$

Se z non cade nell'intervallo (a, b) $\mathcal{R}(z, q)$ è ovviamente limitato come si verifica con la disuguaglianza

$$\int_a^b \left| \frac{f(x)}{z - x} \right|^2 dx \leq \max_{a \leq x \leq b} (|z - x|^{-2}) \|f\|^2$$

Se z appartiene all'intervallo, è necessario restringere l'operatore inverso al dominio di funzioni tali che $f(x)/(z - x)$ sia in L_2 . Inoltre è immediato costruire una successione di vettori f_n con norma uno tali che $\|qf_n\| \rightarrow 0$ ad es.

$$f_n(x) = \begin{cases} \sqrt{n} & |z - x| < \frac{1}{2n} \\ 0 & \text{altrove} \end{cases}$$

Si trova infatti

$$\|(q - z) f_n\|^2 = \int_a^b n(x - z)^2 \theta(2n|x - z| \leq 1) dx \leq \frac{1}{4n}$$

Abbiamo perciò mostrato che $\text{RES}(q) = \{z \in \mathbb{C} \mid z \notin [a, b]\}$, $\mathbf{C}\sigma(q) = [a, b]$, $\mathbf{P}\sigma(q) = \rho(q) = \emptyset$. Si noterà che *il risolvente è una funzione analitica di z con un taglio in corrispondenza dello spettro continuo*; infatti per ogni elemento di matrice del risolvente si ha

$$\langle f, \mathcal{R}(z, q) g \rangle = \int_a^b \frac{\overline{f(x)} g(x)}{z - x} dx$$

Se f e g sono funzioni continue, la discontinuità attraverso il taglio è valutabile facilmente

$$\lim_{\varepsilon \rightarrow 0} \{\langle f, \mathcal{R}(x_0 + i\varepsilon, q) g \rangle - \langle f, \mathcal{R}(x_0 - i\varepsilon, q) g \rangle\} = 2\pi i \overline{f(x_0)} g(x_0).$$

◀

ESEMPIO 4-15. Sia $(\mathcal{D}_*, \mathfrak{C}_*)$ l'operatore definito al § 4.2.5.1. Il risolvente è determinato dalla soluzione dell'equazione

$$((z\mathbb{1} - \mathcal{D}_*) f)(x) = g(x), \quad f \in \mathfrak{C}_*$$

ossia dall'equazione differenziale del primo ordine

$$i f'(x) + z f(x) = g(x)$$

che ha la soluzione generale

$$f(x) = A e^{izx} + \int_a^x e^{iz(x-y)} g(y) dy$$

a cui bisogna imporre la condizione $f \in \mathfrak{C}_*$ ossia $f(a) = f(b) = 0$; ciò implica $A = 0$ e inoltre

$$\int_a^b e^{-izy} g(y) dy = 0$$

Perciò il codominio di $z\mathbb{1} - \mathcal{D}_*$ non è denso ma è costituito dai vettori ortogonali a $f_z(x) = e^{izx} \in L_2$. Si conclude che $\rho(\mathcal{D}_*) = \mathbb{C}$ e $\text{RES}(\mathcal{D}_*) = \mathbf{P}\sigma(\mathcal{D}_*) = \mathbf{C}\sigma(\mathcal{D}_*) = \emptyset$. Sia ora \mathcal{D}_- l'estensione dell'operatore che si ottiene eliminando la condizione $f(b) = 0$. Allora $\mathcal{R}_{z\mathbb{1} - \mathcal{D}_-}$ è denso e il risolvente esiste ed è limitato per ogni z : allora $\text{RES}(\mathcal{D}_-) = \mathbb{C}$ e lo spettro è vuoto (naturalmente lo stesso vale per l'operatore \mathcal{D}_+ che si ottiene scambiando a con b). Infine se scegliamo condizioni periodiche $f(a) = f(b)$ troviamo per il risolvente

$$(\mathcal{R}(z, \mathcal{D}) g)(x) = \frac{e^{izx}}{e^{iz(a-b)} - 1} \int_a^b e^{-izy} g(y) dy + \int_a^x e^{iz(x-y)} g(y) dy$$

a patto che $e^{iz(a-b)} \neq 1$ cioè $z \notin \{2n\pi/(b-a) \mid n \in \mathbb{Z}\}$, che costituisce lo spettro puntuale. ◀

🔗 *Mathematica* può essere utilizzata per esplorare varie condizioni al contorno. Una semplice istruzione permette di ricavare il risolvente per gli operatori differenziali più semplici. Si consideri il codice seguente

```

Resolvent.nb
DSolve[{i f'[x] + z f[x] == g[x], f[0]==f[2Pi] }, f[x], x]
DSolve[{f'[x]+z f[x] == DiracDelta[x], f[-Pi]==f[Pi],
f'[-Pi]==f'[Pi]}, f[x], x]//Simplify
    
```

Dalla prima istruzione si ottiene il risolvente per l'operatore \mathcal{D} e questo presenta poli nei punti $z \in \mathbb{Z}$; la seconda operazione produce il risolvente per l'operatore $-d^2/dx^2$ con condizioni periodiche al contorno, e questo presenta poli semplici per $z = n^2, n \in \mathbb{Z}$, che costituiscono gli autovalori dello spettro puntuale. ◀

4.3.2. Proprietà analitiche del risolvente. Gli elementi di matrice del risolvente, nel caso di spazi finito-dimensionali, sono funzioni razionali, e dunque funzioni analitiche dappertutto eccetto che per un numero finito di poli, che costituiscono lo spettro dell'operatore. Nel caso dello spazio di Hilbert le proprietà analitiche del risolvente non sono così semplici, tuttavia vale ancora il seguente teorema

TEOREMA 4.3.2. *Il risolvente $\mathcal{R}(z, \mathcal{A}) = (z\mathbb{1} - \mathcal{A})^{-1}$ è analitico in z nell'intorno di ogni punto dell'insieme risolvente.*

Teor.4.3.2

PROVA. La serie

$$\sum_0^{\infty} (w - z)^n \mathcal{R}(w, \mathcal{A})^{n+1}$$

converge ad un operatore limitato se $|z - w| \|\mathcal{R}(w, \mathcal{A})\| < 1$. Dunque se $w \in \text{RES}(\mathcal{A})$ e $|w - z| < 1/\|\mathcal{R}(w, \mathcal{A})\|$ la serie definisce l'operatore $\mathcal{R}(z, \mathcal{A})$, il che mostra che $\text{RES}(\mathcal{A})$ è un insieme aperto e che il risolvente è sviluppabile in serie di potenze nell'intorno di ogni punto dell'insieme risolvente. Una dimostrazione più dettagliata fa uso della identità (vedi § 3.6)

$$\mathcal{R}(z, \mathcal{A}) = \mathcal{R}(w, \mathcal{A}) + (w - z) \mathcal{R}(w, \mathcal{A}) \mathcal{R}(z, \mathcal{A})$$

che si può iterare^{††} per ottenere

$$\mathcal{R}(z, \mathcal{A}) = \sum_0^n (z - w)^j \mathcal{R}(w, \mathcal{A})^{-(j+1)} + (w - z)^{n+1} \mathcal{R}(w, \mathcal{A})^{-(n+1)} \mathcal{R}(z, \mathcal{A}).$$

Nell'ipotesi che $|z - w| \|\mathcal{R}(w, \mathcal{A})\| < 1$, $\mathcal{R}(z, \mathcal{A})$ esiste limitato e quindi

$$\left\| \mathcal{R}(z, \mathcal{A}) - \sum_0^n (z - w)^j \mathcal{R}(w, \mathcal{A})^{-(j+1)} \right\| \leq |w - z|^{n+1} \|\mathcal{R}(w, \mathcal{A})^{-1}\|^{n+1} \|\mathcal{R}(z, \mathcal{A})\| \xrightarrow{n \rightarrow \infty} 0.$$

■

$n \neq \infty$ Se \mathcal{A} è illimitato $\text{RES}(\mathcal{A})$ può anche essere vuoto; se invece \mathcal{A} è limitato il teorema precedente è senz'altro operante, però l'insieme $\text{RES}(\mathcal{A})$ non è necessariamente connesso, e quindi non è detto che il risolvente si possa ricostruire dappertutto per continuazione analitica. ◀

TEOREMA 4.3.3. *Lo spettro di un operatore limitato consiste di almeno un punto.*

Teor.4.3.3

PROVA. Assumere il contrario porta ad una contraddizione. Infatti il risolvente è analitico, si annulla per $z \rightarrow \infty$ e nel caso che lo spettro sia vuoto esso risulterebbe regolare in tutto il piano complesso, dunque una costante secondo il teorema di Liouville. Ma ovviamente $(w\mathbb{1} - \mathcal{A})^{-1} = (z\mathbb{1} - \mathcal{A})^{-1}$ è incompatibile con $z \neq w$. ■

🔍 Da notare che un operatore illimitato può avere un risolvente analitico in tutto il piano complesso, come mostra l'esempio dell'operatore \mathcal{D}_- a Pag. 126 (vedi anche il Probl.11). Quello che viene a cadere è il fatto che il risolvente tenda a zero all'infinito; in questi casi è come se lo spettro fosse costituito dal solo punto all'infinito. Fortunatamente questi casi un po' patologici trovano poco spazio nelle applicazioni fisiche. ◀

^{††}Un'equazione del tipo $x = x_0 + Kx$ si "itera" sostituendo alla x del membro di destra dell'equazione con l'intera espressione $x_0 + Kx$ ottenendo così $x = x_0 + K(x_0 + K(x_0 + \dots))$

4.3.3. Gli operatori canonici della Meccanica Quantistica. . In meccanica quantistica si definiscono operatori \mathbf{p} e \mathbf{q} nello spazio di Hilbert detti “operatori canonici”, che obbediscono alle regole di commutazione di Heisenberg (si scelgono unità di misura in cui la costante di Planck valga $\hbar = 1$):

$$\mathbf{q}\mathbf{p} - \mathbf{p}\mathbf{q} = i\mathbb{1}$$

Basandoci su quanto sviluppato fino a questo punto, possiamo dimostrare il seguente

Teor.4.3.4

TEOREMA 4.3.4 (Wintner). *Gli operatori canonici non possono essere entrambi limitati.*

PROVA. Consideriamo i due operatori \mathbf{qp} e \mathbf{pq} entrambi limitati; se lo zero appartiene a $\sigma(\mathbf{q})$ ridefiniamo $\mathbf{q} \mapsto \mathbf{q} + z\mathbb{1}$ che pure soddisfa le regole di commutazione e risulta invertibile a patto di scegliere z sufficientemente grande. Dunque senza perdita di generalità possiamo assumere che \mathbf{q}^{-1} esista limitato. Dalla relazione

$$z\mathbb{1} - \mathbf{qp} = \mathbf{q}(z\mathbb{1} - \mathbf{pq})\mathbf{q}^{-1}$$

segue

$$(z\mathbb{1} - \mathbf{qp})^{-1} = \mathbf{q}(z\mathbb{1} - \mathbf{pq})^{-1}\mathbf{q}^{-1}$$

ovvero \mathbf{pq} e \mathbf{qp} hanno lo stesso spettro non vuoto. Dalla regola di commutazione si ha allora che

$$\sigma(\mathbf{qp}) = \sigma(\mathbf{pq}) = \sigma(\mathbf{qp} - i\mathbb{1})$$

ossia se ne può concludere che se $z \in \sigma(\mathbf{qp})$ allora anche $z - i$ è nello spettro e in generale lo sarà $z - ni$ per n intero qualunque, il che contraddice il fatto che \mathbf{qp} è limitato. ■

🔗 Dal momento che \mathbf{q} e \mathbf{p} devono essere definiti in un opportuno dominio $\mathcal{D}_{\mathbf{q},\mathbf{p}} \subset \mathfrak{H}$, le regole di commutazione sono da intendere più correttamente così

$$\mathbf{qp} - \mathbf{pq} \subset i\mathbb{1}$$

Per maggiori dettagli sul problema della rappresentazione delle regole di commutazione di Heisenberg si veda [Put67] ◀

ESEMPIO 4-16. Gli operatori canonici ammettono una rappresentazione in termini di operatori differenziali

$$(\mathbf{q}f)(x) = xf(x), \quad (\mathbf{p}f)(x) = -i df(x)/dx$$

oppure in termini di matrici infinito-dimensionali

$$\mathbf{p} \mapsto \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \sqrt{1} & 0 & \dots & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \dots & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \dots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \\ \vdots & & & \sqrt{n} & 0 & \sqrt{n+1} \\ & & & & \ddots & \ddots & \ddots \end{pmatrix}, \quad \mathbf{q} \mapsto \frac{1}{i\sqrt{2}} \begin{pmatrix} 0 & \sqrt{1} & 0 & \dots & \dots \\ -\sqrt{1} & 0 & \sqrt{2} & 0 & \dots & \dots \\ 0 & -\sqrt{2} & 0 & \sqrt{3} & 0 & \dots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \\ \vdots & & & -\sqrt{n} & 0 & \sqrt{n+1} \\ & & & & \ddots & \ddots & \ddots \end{pmatrix}$$

Le due rappresentazioni sono equivalenti (E. Schroedinger, 1926). ◀

4.3.4. La teoria spettrale per operatori autoaggiunti. Proprietà fondamentali dello spettro per gli operatori autoaggiunti sono i) l'intero spettro è contenuto sulla retta reale e ii) lo spettro residuo è vuoto. Infatti:

P $\sigma(\mathcal{A})$: da $(\mathcal{A}|z) = z|z)$ segue, come nel caso finito-dimensionale, $\langle z|\mathcal{A}|z\rangle = z\langle z|z\rangle = \overline{\langle z|\mathcal{A}^\dagger|z\rangle}$ e dunque $z \in \mathbb{R}$.

C $\sigma(\mathcal{A})$: $(\mathcal{A} - z\mathbb{1})|z, n\rangle \rightarrow 0 \Rightarrow \Im\{\langle z, n|(\mathcal{A} - z\mathbb{1})|z, n\rangle\} = \Im\{z\langle z, n|z, n\rangle\} \rightarrow 0$, ma $\langle z, n|z, n\rangle = 1$ e dunque $\Im\{z\} = 0$.

$\rho(\mathcal{A})$: è costituito da $z \notin \mathbf{P}\sigma(\mathcal{A})$ tali che $\bar{z} \in \mathbf{P}\sigma(\mathcal{A}^\dagger)$ il che è impossibile se $\mathcal{A} = \mathcal{A}^\dagger$.

Per $\Im\{z\} \neq 0$ il risolvente è un operatore limitato con norma minore o uguale a $|\Im\{z\}|^{-1}$

$$\|(\mathcal{A} - z\mathbb{1})\xi\| \leq \frac{\|\xi\|}{|\Im\{z\}|}$$

(la dimostrazione è lasciata come esercizio). È pure immediato verificare che autovettori appartenenti ad autovalori distinti (nello spettro puntuale) sono ortogonali, esattamente come nel caso finito-dimensionale. La separabilità dello spazio di Hilbert implica allora che $\mathbf{P}\sigma$ è un insieme numerabile.

⚡ Il termine spettro discreto non deve trarre in inganno: seppure $\mathbf{P}\sigma$ sia numerabile, è facile costruire casi in cui $\mathbf{P}\sigma$ riempie densamente un intervallo reale; ad es. sia \mathcal{M} una matrice infinita diagonale che contenga sulla diagonale tutti i razionali contenuti in un intervallo $[a, b]$. In questo caso lo spettro puntuale è denso in $[a, b]$ e lo spettro continuo è dato dai numeri irrazionali in $[a, b]$. Nelle applicazioni alla Meccanica Quantistica tuttavia solo alcuni operatori autoaggiunti si qualificano a rappresentare modelli realistici di Hamiltoniano e molti esempi “patologici” sono solo curiosità matematiche. ◀

4.3.4.1. *Considerazioni euristiche.* A questo punto, per proseguire nello studio delle proprietà generali degli operatori autoaggiunti è necessario introdurre il concetto di *famiglia spettrale*. Procediamo per via euristica basandoci sulla formula di Riesz–Dunford. Consideriamo un caso semplice in cui $\mathbf{P}\sigma(\mathcal{A})$ sia formato da un numero finito di punti z_1, z_2, \dots, z_n e $\mathbf{C}\sigma(\mathcal{A})$ sia un intervallo finito $[a, b]$ sulla retta reale, con $\mathbf{P}\sigma(\mathcal{A}) \cap \mathbf{C}\sigma(\mathcal{A}) = \emptyset$. Abbiamo visto che il risolvente $R(z, \mathcal{A}) = (z\mathbb{1} - \mathcal{A})^{-1}$ ammette uno sviluppo in serie attorno ad ogni punto dell’insieme risolvente e quindi costituisce una funzione analitica con singolarità confinate allo spettro dell’operatore. Si avrà perciò

$$R(z, \mathcal{A}) = \frac{1}{2\pi i} \oint_{\Gamma} \frac{R(w, \mathcal{A})}{w - z} dw$$

essendo Γ un cammino chiuso frontiera di un intorno di $z \notin \sigma(\mathcal{A})$. Deformiamo il contorno in modo che una porzione Γ' sia “gonfiata” e mandata all’infinito del piano complesso (Γ') e un’altra porzione Γ'' contenga al suo interno tutto lo spettro (vedi figura). Come nel caso finito

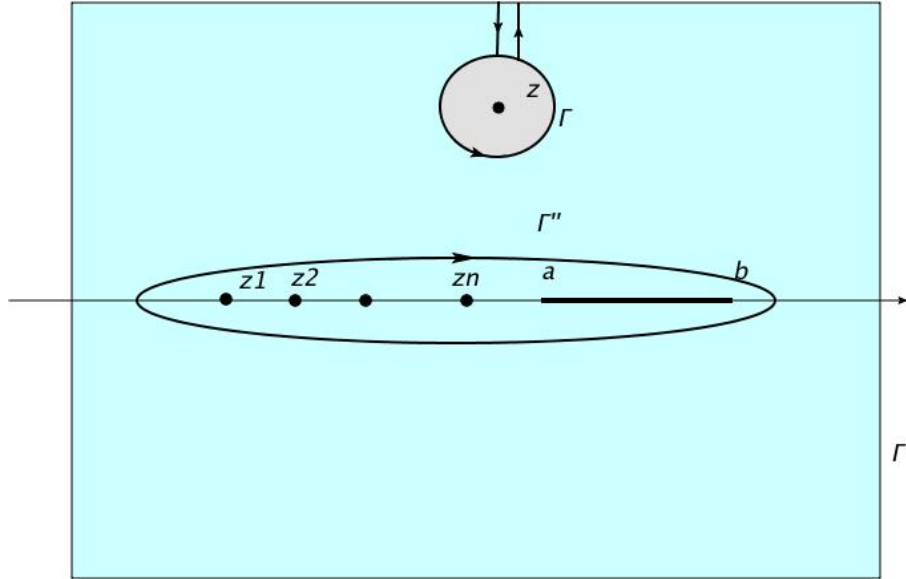


FIGURA 4-2. Il cammino di integrazione Γ viene deformato in modo da adattarsi allo spettro sull’asse reale.

dimensionale l'integrale si interpreta come

$$(4.8) \quad \hat{f}(\mathcal{A}) = \frac{1}{2\pi i} \oint_{\Gamma''} f(w) (z\mathbb{1} - \mathcal{A})^{-1} dw$$

e il cammino Γ'' si può deformare in modo da essere costituito da due rette $\Im m\{w\}=\varepsilon$ e $\Im m\{w\}=-\varepsilon$; si ottiene così la formula

$$\hat{f}(\mathcal{A}) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} [((x - i\varepsilon)\mathbb{1} - \mathcal{A})^{-1} f(x - i\varepsilon) - ((x + i\varepsilon)\mathbb{1} - \mathcal{A})^{-1} f(x + i\varepsilon)] dx$$

che per $\varepsilon \rightarrow 0$ equivale a

$$\hat{f}(\mathcal{A}) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} [((x - i\varepsilon)\mathbb{1} - \mathcal{A})^{-1} - ((x + i\varepsilon)\mathbb{1} - \mathcal{A})^{-1}] f(x) dx$$

e costituisce l'analogo della decomposizione spettrale 3.7 valida per spazi di dimensione finita. Ponendo

$$\mathbb{E}_x = \lim_{\delta \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \frac{1}{2\pi i} \int_{-\infty}^{x+\delta} [((x' - i\varepsilon)\mathbb{1} - \mathcal{A})^{-1} - ((x' + i\varepsilon)\mathbb{1} - \mathcal{A})^{-1}] f(x) dx'$$

la decomposizione spettrale si scrive in modo più compatto

$$\boxed{\hat{f}(\mathcal{A}) = \int_{-\infty}^{\infty} f(x) d\mathbb{E}_x}$$

Si noti che per ogni x reale, \mathbb{E}_x definisce un proiettore. Esaminiamo come si comporta \mathbb{E}_x nell'intorno di un punto dello spettro puntuale. Sia $x_1 < z_1 < x_2$ e si assuma che nell'intervallo (x_1, x_2) non cadano altri punti dello spettro oltre a z_1 . Si trova

$$\begin{aligned} (\mathbb{E}_{x_2} - \mathbb{E}_{x_1}) |z_k\rangle &= \frac{1}{2\pi i} \int_{x_1}^{x_2} [(x - i\varepsilon)\mathbb{1} - \mathcal{A})^{-1} - ((x + i\varepsilon)\mathbb{1} - \mathcal{A})^{-1}] dx |z_k\rangle \\ &= \frac{1}{2\pi i} \oint_{\gamma_1} dz (z - z_k)^{-1} |z_k\rangle = \begin{cases} 0 & k \neq 1 \\ |z_1\rangle & k = 1 \end{cases}. \end{aligned}$$

e quindi, comunque vicini siano x_2, x_1 , $\mathbb{E}_{x_2} - \mathbb{E}_{x_1} = \mathbb{P}_1$ dove \mathbb{P}_1 è il proiettore nel sottospazio appartenente all'autovalore z_1 ; analogamente per tutti gli altri z_k si avrà che \mathbb{E}_x presenta una discontinuità pari a \mathbb{P}_k allorché x passa da sinistra a destra di un punto nello spettro puntuale. Inoltre lo stesso calcolo mostra che $\mathbb{E}_{x'} = \mathbb{E}_{x''}$ qualora nell'intervallo (x', x'') non cadano autovalori di \mathcal{A} . La decomposizione spettrale prende contributo anche dallo spettro continuo; intorno a $z \in \mathbf{C}\sigma$ si ha che \mathbb{E}_x varia con continuità (vedi in seguito il Teor.4.3.6).

4.3.4.2. *Famiglie spettrali.* Le considerazioni euristiche del paragrafo precedente suggeriscono di introdurre la seguente definizione:

Def.4.3.1

DEF. 4.3.1. Una famiglia di operatori autoaggiunti \mathbb{E}_x dipendenti dal parametro reale x si dice una *famiglia spettrale* se

- i) per ogni x reale \mathbb{E}_x è un proiettore ortogonale su di un sottospazio lineare dello spazio di Hilbert;
- ii) per ogni vettore ξ e per $x < y$ si ha

$$\|\mathbb{E}_x \xi\| \leq \|\mathbb{E}_y \xi\|$$

$$\text{iii) } \lim_{x \rightarrow -\infty} \mathbb{E}_x = 0, \quad \lim_{x \rightarrow +\infty} \mathbb{E}_x = \mathbb{1};$$

$$\text{iv) } \lim_{\varepsilon \rightarrow +0} \|(\mathbb{E}_{x+\varepsilon} - \mathbb{E}_x) \xi\| = 0$$

Segue da i) e ii) che $\mathbb{E}_x \mathbb{E}_y = \mathbb{E}_y \mathbb{E}_x = \mathbb{E}_x$ per $x < y$. La condizione iv) è una convenzione che prescrive la *continuità a destra della famiglia spettrale*.

Si possono ora definire integrali sulla famiglia spettrale nel modo seguente: sia $f(x)$ continua in un intervallo (a, b) . Suddividiamo l'intervallo in n parti $x_0 \equiv a, x_1, x_2, \dots, x_{n-1}, x_n \equiv b$ e formiamo le somme parziali

$$S_n = \sum_{k=1}^n f(x_k) (\mathbb{E}_{x_k} - \mathbb{E}_{x_{k-1}})$$

Se $f(x)$ è limitata, S_n definisce un operatore lineare limitato. passiamo ora al limite per $n \rightarrow \infty$ in modo che tutti gli intervalli $x_k - x_{k-1}$ tendano a zero. Il limite si definisce come l'integrale della $f(x)$ rispetto alla famiglia spettrale \mathbb{E}_x e si indica con

$$\int_a^b f(x) d\mathbb{E}_x$$

Si intende che il limite è nel senso della norma dello spazio di Hilbert. È importante la relazione

$$\left\| \int_a^b f(x) d\mathbb{E}_x \xi \right\|^2 = \int_a^b |f(x)|^2 d\langle \xi, \mathbb{E}_x \xi \rangle.$$

valida per ogni ξ nel dominio di \mathcal{A} . Sotto opportune condizioni sulla $f(x)$ si considera anche il caso $a = -\infty$ e $b = \infty$. Enunciamo ora, senza darne dimostrazione, i risultati più importanti della teoria spettrale per gli operatori autoaggiunti nello spazio di Hilbert:

TEOREMA 4.3.5 (I teorema spettrale). *Per ogni operatore autoaggiunto $(\mathcal{A}, D_{\mathcal{A}})$ esiste una ed una sola famiglia spettrale \mathbb{E}_x avente le seguenti caratteristiche*

Teor.4.3.5

- i) $\xi \in D_{\mathcal{A}}$ se-e-solo-se esiste finito l'integrale $\int_{-\infty}^{\infty} x^2 d\langle \xi, \mathbb{E}_x \xi \rangle$
- ii) se $\xi \in D_{\mathcal{A}}$ allora $\mathcal{A}\xi = \int_{-\infty}^{\infty} x d\mathbb{E}_x \xi$ e perciò $\|\mathcal{A}\xi\|^2 = \int_{-\infty}^{\infty} x^2 d\langle \xi, \mathbb{E}_x \xi \rangle$

Viceversa, ogni operatore definito attraverso le condizioni i, ii) a partire da una famiglia spettrale è autoaggiunto. Se \mathcal{A} è limitato, ogni operatore limitato \mathcal{B} che commuti con \mathcal{A} commuta anche con ogni \mathbb{E}_x .

⚡ La presentazione che si è data in questa sezione è molto schematica e si sono esposti solo argomenti di plausibilità per la decomposizione spettrale. Si consiglia perciò di approfondire l'argomento consultando, ad es. [Nai68, Tay58]. ◀

La conoscenza della famiglia spettrale determina completamente lo spettro dell'operatore. Infatti si ha:

TEOREMA 4.3.6 (II Teorema spettrale). *Sia \mathbb{E}_x la famiglia spettrale di un operatore autoaggiunto \mathcal{A} . allora l'asse reale è suddiviso come segue:*

Teor.4.3.6

- a) x_0 è un punto di $\text{RES}(\mathcal{A})$ se \mathbb{E}_x è costante in un intorno di x_0 ;
- b) x_0 è un autovalore ($x_0 \in \text{P}\sigma(\mathcal{A})$) se \mathbb{E}_x è discontinuo in x_0 , cioè $\lim_{h \rightarrow 0+} \mathbb{E}_{x_0-h} \neq \mathbb{E}_{x_0}$;
- c) x_0 appartiene allo spettro continuo se \mathbb{E}_x è continuo in x_0 ma non costante in alcun intorno di x_0 .

PROVA. La dimostrazione è semplice, una volta acquisito il Teorema I. Infatti da

$$\|(\mathcal{A} - x_0 \mathbb{1}) \xi\|^2 = \int_{-\infty}^{\infty} (x - x_0)^2 d\langle \xi, \mathbb{E}_x \xi \rangle$$

segue che se \mathbb{E}_x è costante nell'intervallo $(x_0 - \epsilon, x_0 + \epsilon)$

$$\|(\mathcal{A} - x_0 \mathbb{1}) \xi\|^2 = \int_{|x-x_0|>\epsilon} (x - x_0)^2 d\langle \xi, \mathbb{E}_x \xi \rangle > \epsilon^2 \int_{-\infty}^{\infty} d\langle \xi, \mathbb{E}_x \xi \rangle = \epsilon^2 \|\xi\|^2$$

il che mostra che $x_0 \in \text{res}(\mathcal{A})$. Supponiamo ora che esista il limite $\lim_{\delta \rightarrow 0^+} (\mathbb{E}_{x_0} - \mathbb{E}_{x_0-\delta}) = \mathbb{P} \neq \mathbb{O}$. allora

$$\begin{aligned} \mathcal{A} \mathbb{P} &= \int x d\mathbb{E}_x (\mathbb{E}_{x_0} - \mathbb{E}_{x_0-\delta}) \\ &= \int_{-\infty}^{x_0} x d\mathbb{E}_x - \int_{-\infty}^{x_0-\delta} x d\mathbb{E}_x = \int_{x_0-\delta}^{x_0} x d\mathbb{E}_x \xrightarrow{\delta \rightarrow 0} x_0 \mathbb{P} \end{aligned}$$

e perciò $\mathcal{A} \mathbb{P} \xi = x_0 \mathbb{P} \xi$ per qualche ξ , ossia $x_0 \in \mathbf{P}\sigma(\mathcal{A})$. Se infine \mathbb{E}_x è continuo ma non-costante in un intorno di x_0 si avrà

$$\begin{aligned} \|(\mathcal{A} - x_0 \mathbb{1})(\mathbb{E}_{x_0+\delta} - \mathbb{E}_{x_0-\delta}) \xi\|^2 &= \int_{-\infty}^{\infty} (x - x_0)^2 d\|\mathbb{E}_x(\mathbb{E}_{x_0+\delta} - \mathbb{E}_{x_0-\delta}) \xi\|^2 \\ &= \int_{x_0-\delta}^{x_0+\delta} (x - x_0)^2 d\|\mathbb{E}_x - \mathbb{E}_{x_0-\delta}\| \xi\|^2 \leq \delta^2 \|(\mathbb{E}_{x_0+\delta} - \mathbb{E}_{x_0-\delta}) \xi\|^2 \end{aligned}$$

e perciò $x_0 \in \mathbf{C}\sigma(\mathcal{A})$. ■

⚡ Ci si potrebbe chiedere come si potrebbe caratterizzare il caso di un autovalore dello spettro discreto immerso in un intervallo che rappresenti lo spettro continuo: in tal caso la famiglia spettrale ha una discontinuità e in più non rimane costante in alcun intorno del punto. ◀

ESEMPIO 4-17. Sia q l'operatore in $L_2(a, b)$

$$(qf)(x) = xf(x)$$

Applicando la definizione di \mathbb{E}_x di Pag.4.3.4.1 si trova:

$$\begin{aligned} (\mathbb{E}_t f)(x) &= \lim_{\substack{\delta \rightarrow 0 \\ \epsilon \rightarrow 0}} \int_{-\infty}^{t+\delta} [(t' - i\epsilon - Q)^{-1} - (t' + i\epsilon - Q)^{-1}] f(x) \frac{dt'}{2\pi i} \\ &= \lim_{\substack{\delta \rightarrow 0 \\ \epsilon \rightarrow 0}} \int_{-\infty}^{t+\delta} [(t' - i\epsilon - Qx)^{-1} - (t' + i\epsilon - x)^{-1}] f(x) \frac{dt'}{2\pi i} \\ &= \lim_{\substack{\delta \rightarrow 0 \\ \epsilon \rightarrow 0}} \int_{-\infty}^{t+\delta} dt' \frac{\epsilon/\pi}{(t' - x)^2 + \epsilon^2} f(x) \\ &= \begin{cases} 0 & t < x \\ f(x) & t \geq x \end{cases} \end{aligned}$$

ossia

$$(\mathbb{E}_t f)(x) = \theta(x \leq t) f(x)$$

ESEMPIO 4-18. L'operatore $p = -i \frac{d}{dx}$ in $L_2(-\infty, +\infty)$: mostreremo nel §4.3.4.3 che il risolvente di p è dato da

$$((k \mp i\epsilon - p)^{-1} f)(x) = i \int_x^{\pm\infty} e^{i(k \mp i\epsilon)(x-y)} f(y) dy$$

e quindi, passando al limite su δ, ϵ ,

$$\begin{aligned} (\mathbb{E}_k f)(x) &= \frac{1}{2\pi} \int_x^{k+\delta} ((k' - i\epsilon - p)^{-1} - (k' + i\epsilon - p)^{-1}) dk' f(x) \\ &= \frac{1}{2\pi} \int_{-\infty}^{k+\delta} dk' \int_{-\infty}^{+\infty} e^{ik'(x-y)} e^{-\epsilon|x-y|} f(y) dy \end{aligned}$$

⚡ Si noterà che $\mathbb{E}_\infty = \mathbb{1}$ equivale alla formula di inversione di Fourier. Il formalismo introduce spontaneamente il fattore di “smorzamento adiabatico” $\exp\{-\epsilon|x-y|\}$ che permette di trattare qualunque f in L_2 , mentre la formula di Fourier senza regolatore si applica solo a $L_2 \cap L_1$. ◀

⚡ L'uso della famiglia spettrale \mathbb{E}_x è conveniente per descrivere lo spettro continuo, mentre nei punti dello spettro puntuale è più semplice e naturale passare ai proiettori \mathbb{P} dati dalla discontinuità di \mathbb{E}_x . La decomposizione spettrale è dunque scritta più convenientemente nella forma

$$\hat{f}(\mathcal{A}) = \sum_{x \in \text{P}\sigma(\mathcal{A})} f(x) \mathbb{P}_x + \int_{x \in \text{C}\sigma(\mathcal{A})} f(x) d\mathbb{E}_x$$

Nei testi di Meccanica Quantistica si suole porre formalmente

$$d\mathbb{E}_x := |x\rangle dx \langle x|, \quad \text{con } \langle x|x'\rangle = \delta(x-x').$$

Questo formalismo (alla Dirac) è euristicamente efficace, ma bisogna tenere presente che una trattazione rigorosa in questi termini richiederebbe l'estensione dello spazio di Hilbert ad uno spazio di distribuzioni ([Nai68], §21). ◀

4.3.4.3. *Operatori differenziali in $L_2(-\infty, \infty)$.* Sia $\mathbf{p} = -i\hbar/dx$ l'operatore già considerato nel capitolo precedente a Pag. 132. Il dominio sia definito da

$$D_{\mathbf{p}} = \{f \in L_2(-\infty, \infty), \text{ a.c.}, \frac{df}{dx} \in L_2\}$$

L'operatore \mathbf{p} è autoaggiunto. Vogliamo costruire l'operatore risolvante di \mathbf{p} . Si tratta di risolvere l'equazione differenziale

$$((k\mathbb{1} - \mathbf{p})f)(x) = g(x)$$

per $\Im\{k\} \neq 0$. La soluzione generale è ovviamente

$$f(x) = -i \int_a^x e^{ik(x-y)} g(y) dy$$

con a reale arbitraria. Si tratta di fissare a in modo che $f \in L_2$ per qualunque $g \in L_2$. Dato che $f(x)$ è esprimibile come convoluzione

$$f(x) = i \int_0^{x-a} e^{ikt} g(x-t) dt$$

una disuguaglianza dovuta a Young (vedi commento a Pag. 134) permette di concludere che deve porsi

$$\begin{cases} a = -\infty & \text{per } \Im\{k\} > 0 \\ a = +\infty & \text{per } \Im\{k\} < 0 \end{cases}$$

infatti

$$\begin{cases} e^{ikt}\theta(t) \in L_1(-\infty, \infty) & \text{per } \Im\{k\} > 0 \\ e^{ikt}\theta(-t) \in L_1(-\infty, \infty) & \text{per } \Im\{k\} < 0 \end{cases}$$

Si ha perciò

$$(R_k(\mathbf{p})f)(x) = ((k\mathbb{1} - \mathbf{p})^{-1}f)(x) = \begin{cases} -i \int_{-\infty}^x e^{ik(x-y)} f(y) dy & \text{per } \Im\{k\} > 0 \\ i \int_x^{\infty} e^{ik(x-y)} f(y) dy & \text{per } \Im\{k\} < 0 \end{cases}$$

Si veda anche il capitolo successivo 4.3.7. Si trova perciò che il risolvante di \mathbf{p} è una funzione analitica per $\Im\{k\} \neq 0$ con una discontinuità $R_{k+i\epsilon}(\mathbf{p}) - R_{k-i\epsilon}(\mathbf{p}) \neq 0$ su tutto l'asse reale che costituisce lo spettro continuo di \mathbf{p} . Si può studiare in modo analogo l'operatore $-\frac{d^2}{dx^2}$ che risulta

autoaggiunto nel dominio costituito dalle *funzioni a.c. con derivata a.c. e derivata seconda in L_2* . Il risolvente di p^2 si può ottenere da quello appena calcolato essendo

$$(k^2 \mathbb{1} - p^2)^{-1} = \frac{1}{2k} ((k\mathbb{1} - p)^{-1} + (k\mathbb{1} + p)^{-1})$$

Si trova facilmente

$$((k^2 \mathbb{1} - p^2)^{-1} f)(x) = \frac{1}{2ki} \int_{-\infty}^{\infty} e^{ik|x-y|} f(y) dy$$

Nella variabile complessa $\lambda = k^2$ il risolvente presenta un taglio sull'asse reale positivo che rappresenta lo spettro continuo di p^2 .

Facendo uso dell'espressione esplicita del risolvente possiamo calcolare *qualunque* funzione di p^2 , ad es.

$$\begin{aligned} (e^{itp^2/2} f)(x) &= \frac{1}{2\pi i} \int_{\Gamma} (\lambda \mathbb{1} - p^2)^{-1} e^{it\lambda/2} f(x) d\lambda \\ &= -\frac{1}{4\pi} \int_{\Gamma} \frac{e^{it\lambda/2}}{\sqrt{\lambda}} \int_{-\infty}^{\infty} e^{i\sqrt{\lambda}|x-y|} f(y) dy \end{aligned}$$

dove Γ è un cammino che abbraccia lo spettro di p^2 come in Fig.4-3.

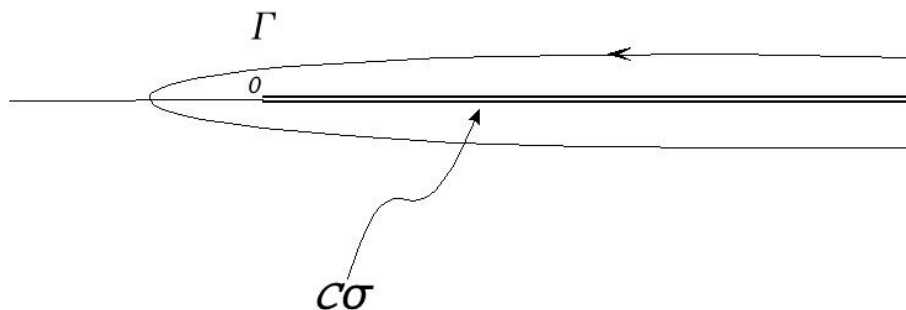


FIGURA 4-3. Il cammino di integrazione Γ nel calcolo di $f(p^2)$.

Si trova pertanto

$$(e^{itp^2/2} f)(x) = -\frac{1}{4\pi} \int_{-\infty}^{\infty} dy f(y) \int_{\Gamma} \frac{d\lambda}{\sqrt{\lambda}} e^{it\lambda/2 + i\sqrt{\lambda}|x-y|}$$

e attraverso un cambiamento di variabile $\lambda = k^2$ si ottiene

$$\dots = \frac{1}{2\pi} \int_{-\infty}^{\infty} dy f(y) \int_{-\infty}^{\infty} dk \exp\{it k^2/2 + ik|x-y|\}$$

L'integrale gaussiano si calcola immediatamente per dare il risultato finale

$$(e^{itp^2/2} f)(x) = \frac{1}{\sqrt{2i\pi t}} \int_{-\infty}^{\infty} dy f(y) e^{\frac{(x-y)^2}{2it}}$$

Naturalmente lo stesso risultato si ottiene facilmente utilizzando direttamente la trasformata integrale di Fourier.

⚡ *Disuguaglianza di Young: sia*

$$(f \star g)(x) = \int_{-\infty}^{\infty} f(y) g(x-y) dy$$

allora $\|f \star g\|_2 \leq \|f\|_1 \|g\|_2$, essendo $\|f\|_q = (\int_{-\infty}^{\infty} |f(x)|^q dx)^{1/q}$. Si veda [RS78] per una disuguaglianza più generale e per la dimostrazione. ◀

4.3.5. Operatori compatti. Una classe di operatori che presenta interessanti proprietà spettrali è quella degli operatori compatti (detti anche “completamente continui”). Premettiamo la seguente definizione:

DEF. 4.3.2. un sottoinsieme M in uno spazio unitario completo (finito o infinito dimensionale) è detto compatto se ogni successione infinita ξ_n contenuta in M ha almeno un punto di accumulazione in M e cioè da ogni successione è possibile estrarre una sottosuccessione convergente in M .

Def.4.3.2

$n \neq \infty$ Ricordiamo che ogni sottoinsieme chiuso e limitato di uno spazio a n dimensioni (reale o complesso) è compatto (teorema di Bolzano–Weierstrass). Ciò non è più vero in uno spazio di Hilbert. Ad es. la sfera unitaria $S^\infty = \{\xi \in \mathfrak{H} \mid \|\xi\| = 1\}$ è un insieme chiuso, limitato ma non compatto; in effetti ogni successione di vettori ortonormali ξ_n ($\langle \xi_n | \xi_m \rangle = \delta_{nm}$) è tale che $\|\xi_i - \xi_j\| = \sqrt{2}$ e quindi non ha alcun punto di accumulazione. ◀

Proprietà importante degli insiemi compatti è quella relativa alla possibilità di ricoprirli con un numero *finito* di aperti; più precisamente sia B_n una qualunque successione di aperti tali che l'unione $\bigcup_n B_n$ contenga l'insieme compatto M . Allora esiste sempre un numero finito di aperti $B_{n_1}, B_{n_2}, \dots, B_{n_k}$ la cui unione contiene M (“da ogni ricoprimento si può estrarre un ricoprimento finito”). Daremo qui di seguito una lista di definizioni e di teoremi senza dimostrazioni. Si vedano [Nai68, Tay58, RS80] per maggiori dettagli.

DEF. 4.3.3. Un operatore lineare \mathcal{A} nello spazio di Hilbert \mathfrak{H} è detto *compatto* se esso trasforma ogni sottoinsieme chiuso e limitato in \mathfrak{H} in un sottoinsieme compatto.

Def.4.3.3

TEOREMA 4.3.7. Ogni operatore compatto è limitato.

Teor.4.3.7

TEOREMA 4.3.8. Se \mathcal{A} è compatto e \mathcal{B} è continuo allora $\mathcal{A}\mathcal{B}$ e $\mathcal{B}\mathcal{A}$ sono compatti.

Teor.4.3.8

TEOREMA 4.3.9. Se \mathcal{A}_n è una successione di operatori compatti ed esiste un operatore \mathcal{A} tale che

Teor.4.3.9

$$\forall \xi \in \mathfrak{H} : \quad \|(\mathcal{A} - \mathcal{A}_n)\xi\| < \frac{1}{n} \|\xi\|$$

allora anche \mathcal{A} è compatto.

TEOREMA 4.3.10. Sia $\mathcal{A} \leftrightarrow \|a_{ij}\|$ la matrice rappresentativa di \mathcal{A} in una base ortonormale. Se

Teor.4.3.10

$$\sum_{i,j=1}^{\infty} |a_{ij}|^2 = \text{Tr}(\mathcal{A}^\dagger \mathcal{A}) < \infty$$

allora \mathcal{A} è compatto.

TEOREMA 4.3.11. Sia \mathcal{K} l'operatore in $L_2(a, b)$ definito da

Teor.4.3.11

$$(\mathcal{K}f)(x) = \int_a^b K(x, y) f(y) dy.$$

Se

$$(4.9) \quad \int_a^b dx \int_a^b dy |K(x, y)|^2 < \infty$$

allora \mathcal{K} è un operatore compatto.

⚡ Le condizioni dei due teoremi precedenti sono solo sufficienti per la compattezza. Si parla in questi casi di operatori di Hilbert–Schmidt che costituiscono una sottoclasse di operatori compatti. ◀

Teor.4.3.12

TEOREMA 4.3.12 (Teorema spettrale per operatori compatti). *Lo spettro di un operatore compatto è costituito da un insieme finito o infinito di punti aventi come unico punto di accumulazione lo zero. Ogni punto dello spettro $\neq 0$ è un autovalore discreto ($\in \mathbf{P}\sigma$) con molteplicità finita; lo zero può appartenere a $\mathbf{C}\sigma$ oppure a ρ .*

Come si nota, ad eccezione di quanto può accadere nello zero, gli operatori compatti presentano proprietà spettrali molto simili agli operatori finito-dimensionali. Una proprietà di frequente applicabilità è la seguente:

Teor.4.3.13

TEOREMA 4.3.13. *Sia \mathcal{A} un operatore simmetrico tale che il suo risolvente esista e sia un operatore compatto per almeno un punto $z \in \text{RES}(\mathcal{A})$. Allora lo spettro di \mathcal{A} consiste di soli autovalori isolati con molteplicità finita che si accumulano all'infinito. Inoltre il risolvente è compatto per tutti i punti dell'insieme risolvente.*

La teoria degli operatori compatti si è sviluppata in connessione con lo studio delle equazioni integrali. Sia $g \in L_2(a, b)$ e si cerchi la soluzione dell'equazione

$$z f(x) - \int_a^b K(x, y) f(y) dy = g(x)$$

dove $z \notin \sigma(K)$, il “nucleo” K sia simmetrico ($K(x, y) = K(y, x)$) e soddisfi la Eq.4.9; per i teoremi enunciati in precedenza si avrà

$$K(x, y) = \sum_{n=1}^{\infty} \kappa_n u_n(x) \overline{u_n(y)}$$

con $\kappa_n \rightarrow 0$ per $n \rightarrow \infty$. La soluzione è allora

$$f(x) = \sum_{n=1}^{\infty} (z - \kappa_n)^{-1} u_n(x) \int_a^b \overline{u_n(y)} g(y) dy$$

(si vedano [Tay58, RS78, CH62]).

4.3.6. L'estensione di operatori simmetrici. Accenniamo ora al problema di determinare se un operatore simmetrico ammetta o no un'estensione autoaggiunta. Abbiamo visto che se $(\mathcal{A}, D_{\mathcal{A}})$ è un operatore simmetrico e se $(\mathcal{B}, D_{\mathcal{B}})$ è un'estensione simmetrica di \mathcal{A} vale la relazione

$$\mathcal{A} \subset \mathcal{B} \subseteq \mathcal{B}^{\dagger} \subseteq \mathcal{A}^{\dagger}$$

e quindi in ogni caso un'estensione simmetrica di \mathcal{A} è una restrizione di \mathcal{A}^{\dagger} . Un operatore simmetrico è detto *massimale* se esso non ammette alcuna estensione simmetrica. Un operatore autoaggiunto è in particolare massimale.

Sia \mathcal{A} un operatore simmetrico e z un numero complesso con $\Im\{z\} \neq 0$. Indichiamo con R_+ e R_- rispettivamente il codominio di $\mathcal{A} - i\mathbb{1}$ e quello di $\mathcal{A} + i\mathbb{1}$:

$$R_+ = (\mathcal{A} - i\mathbb{1}) D_{\mathcal{A}}, \quad R_- = (\mathcal{A} + i\mathbb{1}) D_{\mathcal{A}}.$$

Si chiamano *spazi di difetto* di \mathcal{A} i sottospazi \mathcal{N}_+ e \mathcal{N}_- che sono rispettivamente i complementi ortogonali di R_+ e di R_- . E' chiaro che \mathcal{N}_+ e \mathcal{N}_- sono individuati dagli autovettori di \mathcal{A}^{\dagger} appartenenti a i e $-i$ rispettivamente (utilizzando lo stesso argomento già visto nella definizione dello spettro residuo)*.

L'operatore

$$\mathcal{V} = (\mathcal{A} - i\mathbb{1}) (\mathcal{A} + i\mathbb{1})^{-1}$$

è detta la *trasformata di Cayley* di \mathcal{A} ed è caratterizzata dalle proprietà seguenti:

- (a) \mathcal{V} è un operatore isometrico con dominio R_- e codominio R_+ .

*La scelta di i è convenzionale, infatti è del tutto equivalente scegliere un qualunque altro numero complesso con parte immaginaria positiva.

- (b) L'insieme dei vettori $\{\mathcal{V}\xi - \xi \mid \xi \in \mathcal{R}_-\}$ è denso in \mathfrak{H} .
(c) Ogni operatore isometrico che soddisfi le condizioni precedenti è la trasformata da Cayley di un operatore simmetrico.

È possibile ora lavorare su \mathcal{V} anziché su \mathcal{A} , con il vantaggio di maneggiare operatori limitati. Ogni estensione isometrica di \mathcal{V} che soddisfi ai requisiti precedenti sarà associata ad una estensione simmetrica di \mathcal{A} .

TEOREMA 4.3.14. *Se \mathcal{A} è chiuso e simmetrico, allora $\mathcal{D}_{\mathcal{A}}$, \mathcal{N}_+ e \mathcal{N}_- sono linearmente indipendenti e la loro somma diretta coincide con $\mathcal{D}_{\mathcal{A}^\dagger}$, cioè ogni vettore $\xi \in \mathcal{D}_{\mathcal{A}^\dagger}$ è decomponibile in modo unico nella somma*

Teor.4.3.14

$$\xi = \xi_0 + \eta_+ + \eta_-, \quad \begin{cases} \xi_0 \in \mathcal{D}_{\mathcal{A}} \\ \eta_{\pm} \in \mathcal{N}_{\pm} \end{cases}$$

Questa decomposizione fornisce una rappresentazione completa di \mathcal{A}^\dagger nella sua azione in $\mathcal{D}_{\mathcal{A}^\dagger}$:

$$\mathcal{A}^\dagger \xi = \mathcal{A} \xi_0 - i\eta_+ + i\eta_-.$$



TEOREMA 4.3.15. *Siano $m_{\pm} = \dim \mathcal{N}_{\pm}$; m_+ e m_- vengono chiamati gli indici di difetto dell'operatore simmetrico \mathcal{A} .*

Teor.4.3.15

TEOREMA 4.3.16. *Un operatore simmetrico ammette almeno un'estensione autoaggiunta se e solo se i suoi indici di difetto sono uguali (se sono entrambi nulli, l'operatore è essenzialmente autoaggiunto; se gli indici sono uguali e positivi, l'operatore ammette infinite estensioni autoaggiunte).*

Teor.4.3.16

Nel caso di indici di difetto uguali, esiste un procedimento costruttivo per determinare l'estensione autoaggiunta dell'operatore simmetrico. Si sceglie una isometria \mathcal{U} da \mathcal{N}_+ a \mathcal{N}_- , e si pone

$$\mathcal{D}_{\mathcal{B}} = \{\xi + \eta_+ + \mathcal{U}\eta_+ \mid \xi \in \mathcal{D}_{\mathcal{A}}, \eta_+ \in \mathcal{N}_+\}$$

$$\mathcal{B}(\xi + \eta_+ + \mathcal{U}\eta_+) = \mathcal{A}\xi + i\eta_+ - i\mathcal{U}\eta_+$$

Si veda [RS78]. Le estensioni autoaggiunte sono dunque in corrispondenza con gli elementi del gruppo $\mathcal{U}(m_+)$ con m_+ parametri reali.



ESEMPIO 4-19. Il caso più semplice è quello dell'operatore \mathcal{D}_0 di Pag. 120 che definiamo per semplicità nell'intervallo simmetrico $(-\log(2), \log(2))$. Gli spazi di difetto sono caratterizzati da

$$-i \frac{d}{dx} f_{\pm}(x) = \pm i f_{\pm}(x) \Rightarrow f_{\pm}(x) = C e^{\mp x}$$

e le isometrie da \mathcal{N}_- a \mathcal{N}_+ sono parametrizzate da una fase γ :

$$\mathcal{U}: f_- \Rightarrow e^{i\gamma} f_+$$

In conclusione il dominio della estensione autoaggiunta \mathcal{D} è dato da

$$\mathcal{D} = \{f_0 + C e^{-x} + C e^{i\gamma} e^x \mid f_0 \in \mathcal{D}_0\}$$

e l'azione di \mathcal{D} è semplicemente

$$\mathcal{D}(f_0 + C e^{-x} + C e^{i\gamma} e^x) = f_0 + i C e^{-x} - i C e^{i\gamma} e^x$$

Si noti che le funzioni in \mathcal{D} soddisfano la condizione al contorno

$$f(\log(2)) = C/2 + 2e^{i\gamma}, \quad f(-\log(2)) = 2C + e^{i\gamma}/2$$

e quindi il rapporto

$$f(\log(2))/f(-\log(2)) = \frac{\frac{1}{2} + 2e^{i\gamma}}{2 + \frac{1}{2}e^{i\gamma}}$$

risulta di modulo uno, il che corrisponde a quanto già visto in precedenza. ◀

4.3.7. Teorema di Stone. abbiamo visto nella Parte II lo stretto legame che intercorre tra gli operatori autoaggiunti e gli operatori unitari. Questo legame è importante nelle applicazioni alla Meccanica Quantistica per la costruzione della dinamica a partire dall'operatore Hamiltoniano. In uno spazio di Hilbert vale il seguente risultato generale:

TEOREMA DI STONE, 1932. *Sia $\mathcal{U}(t)$ una famiglia di operatori unitari, parametrizzata dalla variabile reale t , che soddisfi alle seguenti condizioni*

- i) $\mathcal{U}(0) = \mathbb{1}$
- ii) $\mathcal{U}(t_1)\mathcal{U}(t_2) = \mathcal{U}(t_1 + t_2)$
- iii) $\mathcal{U}(t)$ è continuo in t nel senso che $\langle \xi, \mathcal{U}(t)\eta \rangle$ è una funzione continua per ogni scelta di vettori $\xi, \eta \in \mathfrak{H}$.

Allora esiste un unico operatore autoaggiunto $(\mathcal{A}, D_{\mathcal{A}})$ tale che

- *) $\mathcal{U}(t) = \exp\{it\mathcal{A}\}$
- **) $D_{\mathcal{A}} = \{\xi \in \mathfrak{H} | \exists \lim_{t \rightarrow 0} t^{-1}(\mathcal{U}(t) - \mathbb{1})\xi\}$

L'operatore \mathcal{A} è detto il *generatore infinitesimale* del gruppo a un parametro $\mathcal{U}(t)$.

Sia ad esempio $\mathcal{T}(a)$ il gruppo di trasformazioni unitarie in $L_2(\mathbb{R})$ definito a Pag. 118

$$(\mathcal{T}(a)f)(x) = f(x - a)$$

L'operatore $\mathcal{T}(a)$ è unitario per ogni valore di a , soddisfa alle condizioni del teorema di Stone: in particolare la continuità

$$\langle f, \mathcal{T}(a)g \rangle = \int_{\mathbb{R}} dx \overline{f(x)} g(x - a) \xrightarrow{a \rightarrow 0} 0$$

discende dalle proprietà della convoluzione. Il corrispondente generatore infinitesimale è proprio l'operatore differenziale $\mathfrak{p} = -i \frac{d}{dx}$ che rappresenta il momento lineare in Meccanica Quantistica.

⚡ Si distinguono vari tipi di continuità nel contesto dei gruppi di operatori. Quella che abbiamo inserito nell'enunciato del teorema di Stone è detta *continuità debole*. Per *continuità forte* si intende che

$$\|(\mathcal{U}(a) - \mathbb{1})\xi\| \xrightarrow{a \rightarrow 0} 0$$

per ogni $\xi \in \mathfrak{H}$. Per *continuità in norma* infine si intende che

$$\|\mathcal{U}(a) - \mathbb{1}\| \xrightarrow{a \rightarrow 0} 0$$

Quest'ultima **non** è necessaria per il teorema di Stone, ad es. non è realizzata nel caso di $\mathcal{T}(.)$; infatti per ogni a si può trovare uno stato η_a tale che $\mathcal{T}(a)\eta_a$ sia ortogonale a η_a . D'altro canto dalla relazione

$$\|(\mathcal{U}(a) - \mathbb{1})\xi\|^2 = \|\mathcal{U}\xi\|^2 - \langle \xi, \mathcal{U}(a)\xi \rangle - \langle \mathcal{U}(a)\xi, \xi \rangle + \|\xi\|^2 \xrightarrow{a \rightarrow 0} 0$$

si vede che nel caso di gruppi di operatori unitari la *continuità debole* implica quella *forte*. Si veda [RS80] per una trattazione più ampia dell'argomento. ◀

Si rivelano spesso utili le due formule seguenti che legano il risolvente del generatore infinitesimale alla trasformata di Laplace di $\mathcal{U}(t)$:

$$(z\mathbb{1} - \mathcal{A})^{-1} = -i \int_0^{\infty} e^{izt} e^{-it\mathcal{A}} dt \quad (\Im\{z\} > 0)$$

$$(z\mathbb{1} - \mathcal{A})^{-1} = i \int_{-\infty}^0 e^{izt} e^{-it\mathcal{A}} dt \quad (\Im\{z\} < 0)$$

valida anche nella forma reale

$$(4.10) \quad (\mathcal{A} - z\mathbb{1})^{-1} = \int_0^{\infty} e^{-t(\mathcal{A} - z\mathbb{1})} dt$$

4.3.8. L'integrale di Fourier. Consideriamo lo spazio di Hilbert $L_2(\mathbb{R})$; si definisce “trasformata (integrale) di Fourier” l'operatore lineare

$$(4.11) \quad (\mathcal{F}f)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixy} f(y) dy$$

definito sulle funzioni $\{f \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})\}$. Le proprietà principali di \mathcal{F} sono le seguenti:

- i) $(\mathcal{F}f)(x)$ è continua e tende a zero per $|x| \rightarrow \infty$ (Lemma di Riemann–Lebesgue)
- ii) Se $f \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$ è anche continua, allora vale la formula inversa

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ixy} (\mathcal{F}f)(y) dy$$

- iii) \mathcal{F} è un operatore isometrico ($\|\mathcal{F}f\| = \|f\|$);
- iv) \mathcal{F} è un operatore unitario (infatti il codominio di \mathcal{F} contiene tutte le funzioni continue di $L_1 \cap L_2$ ed è quindi denso in L_2);
- v) \mathcal{F} si estende per continuità a un operatore unitario in $L_2(\mathbb{R})$; la forma esplicita della trasformata, valida per ogni vettore in L_2 è data da

$$(\mathcal{F}f)(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} \int_{-\infty}^{\infty} \frac{e^{ixy} - 1}{iy} f(y) dy$$

(la derivazione sotto segno di integrale essendo permessa solo se $f \in D_{\mathcal{F}}$)[†]

 In Meccanica Quantistica la trasformata di Fourier effettua una trasformazione di similitudine dall'operatore \mathfrak{p} all'operatore \mathfrak{q} :

$$(\mathfrak{p}f)(x) = -i \frac{d}{dx} f(x) = \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} e^{ixy} y (\mathcal{F}^{-1}f)(y) = (\mathcal{F}\mathfrak{q}\mathcal{F}^{\dagger}f)(x)$$

In conseguenza di questo fatto si ha anche $\mathcal{F}(\mathfrak{p}^2 + \mathfrak{q}^2) = (\mathfrak{p}^2 + \mathfrak{q}^2)\mathcal{F}$ e dato che l'operatore $\mathfrak{p}^2 + \mathfrak{q}^2$ è autoaggiunto con tutti autovalori semplici (si veda un qualunque testo di Meccanica Quantistica) \mathcal{F} è una funzione di $\mathfrak{p}^2 + \mathfrak{q}^2$. Si verifica poi facilmente che vale l'identità

$$\mathcal{F} = \exp\{-i\frac{\pi}{4}(\mathfrak{p}^2 + \mathfrak{q}^2 - \mathbb{1})\}$$

Questo risultato è collegato a proprietà dei polinomi di Hermite sotto trasformata di Fourier (vedi il § esercizi). 

[†]Alternativamente si può definire l'integrale come “limite-in-media” $\lim_{M \rightarrow \infty} \frac{1}{M} \int_{-M}^{+M} \exp\{ixy\} f(y) dy$

Appendice III

4.4. La formula di Lie–Trotter e la Meccanica Quantistica

La formula di Lie 3.4.3.1 si estende sotto opportune condizioni al caso di operatori nello spazio di Hilbert (Trotter) [Nel67]. Si è riconosciuto in tempi abbastanza recenti che questa formula è alla base dell’approccio di Feynman in termini di “path integrals”. Al di là degli aspetti formali, di per sè molto interessanti, la formula di Lie–Trotter e le sue modifiche tecniche si presta a realizzare algoritmi di calcolo molto efficaci per la soluzione dell’equazione di Schroedinger in tempo reale oppure alla formulazione di algoritmi Monte Carlo in tempo euclideo.

Consideriamo l’operatore Hamiltoniano che descrive una singola particella scalare di massa m immersa in un campo di forze conservativo con energia potenziale $V(x)$ (ci limitiamo al caso di un grado di libertà, ma l’estensione a più gradi di libertà non offre difficoltà, come pure al caso di energia potenziale dipendente dal tempo):

$$\mathcal{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \equiv \mathcal{K} + \mathcal{V}$$

La soluzione dell’equazione di Schroedinger, utilizzando per semplicità di notazione unità in cui $\hbar = 1$, è data formalmente da

$$\psi(t) = \exp\{-i\mathcal{H}t\} \psi(0)$$

Applicando la formula di Lie otteniamo

$$(4.12) \quad \exp\{-i\mathcal{H}t\} = \lim_{n \rightarrow \infty} (\exp\{-it\mathcal{K}/n\} \exp\{-it\mathcal{V}/n\})^n$$

Considerati separatamente \mathcal{K} e \mathcal{V} si esponenziano facilmente, il che rende conveniente l’uso della formula (detta anche “splitting method”):

$$(4.13) \quad \begin{aligned} (\exp\{-it\mathcal{K}\}\psi)(x) &= \sqrt{\frac{m}{2\pi it}} \int_{\mathbb{R}} \exp\left\{\frac{im(x-y)^2}{2t}\right\} \psi(y) dt \\ (\exp\{-it\mathcal{V}\}\psi)(x) &= \exp\{-itV(x)\} \psi(x) \end{aligned}$$

Si ottiene così

$$\begin{aligned} &(\exp\{-it\mathcal{K}/n\} \exp\{-it\mathcal{V}/n\})^n \psi(x) = \\ &\left(\frac{2\pi it}{m}\right)^{-n/2} \int \dots \int dx_1 \dots dx_n \exp\left\{i\frac{m}{2} \sum_{i=1}^n \frac{(x_i - x_{i+1})^2}{t/n} - \frac{it}{n} \sum_{i=1}^n V(x_i)\right\} \psi(x_1) \end{aligned}$$

La fase all’esponente è interpretabile come una approssimazione discreta della azione classica

$$\frac{t}{n} \sum_{i=1}^n \left\{ \frac{m}{2} \left(\frac{x_i - x_{i+1}}{t/n} \right)^2 - V(x_i) \right\}$$

Se ignoriamo il fatto che tutte le variabili x_i vanno integrate su \mathbb{R} e quindi gli incrementi $x_{i+1} - x_i$ non sono in alcun senso infinitesimali, possiamo identificare questa somma con l’approssimante N -esima dell’integrale

$$\int_0^t \left[\frac{1}{2} m \left(\frac{dx(\tau)}{d\tau} \right)^2 - V(x(\tau)) \right] d\tau$$

Si arriva così all'espressione del propagatore quantistico in termini di “somma sui cammini” (Feynman, 1946) [FH65].

4.5. Metodi computazionali

La disponibilità di calcolatori veloci è ormai diffusa ad ogni livello e molti problemi che hanno richiesto analisi complicate in passato ora sono alla portata di tutti con minimo sforzo. Basti pensare allo studio della meccanica celeste, un'impresa che ha impegnato fisici e astronomi per più di un secolo nello studio della dinamica del sistema solare mentre oggi rappresenta, nella sua forma più elementare, poco più di un esercizio scolastico. Ci occuperemo qui di due tipi di problemi: evoluzione temporale di sistemi lineari, avendo in mente tipicamente l'equazione di Schroedinger, e lo studio dello spettro puntuale. In entrambi i casi è necessario un primo passo che consiste nel ridurre la dimensione infinita dello spazio degli stati (lo spazio di Hilbert) a una dimensione grande ma finita N , rappresentabile nella memoria di un computer. È necessario a questo scopo introdurre una “scatola di normalizzazione” (o cutoff infrarosso) caratterizzata da una taglia di lunghezza L . I parametri N ed L entrano come parametri “tecnici” nel problema fisico accanto ad altri parametri che invece sono dettati dalla fisica del problema. La strategia di calcolo numerico deve ottimizzare la scelta di N ed L in modo da ridurre l'errore introdotto dalla loro presenza senza portare a livelli insopportabili lo spazio in memoria necessario e il tempo di calcolo. La regola base è, detta sbrigativamente, la seguente: se la scala fisica del problema è ℓ (ad es. il raggio di Bohr in fisica atomica, o il fermi in fisica nucleare) il rapporto ℓ/L deve essere tenuto il più piccolo possibile, tenendo conto che l'errore coinvolto può essere del tipo $O(\exp\{-\kappa L/\ell\})$. D'altra parte se la scala di energia da esplorare è dell'ordine Λ , detto $\alpha = L/N$ (il “passo reticolare”) è necessario controllare che sia rispettata una relazione del tipo $\alpha^\nu \Lambda \ll 1$ dove ν dipende dalla natura del problema - ad es. per l'equazione di Schroedinger non-relativistica $\nu = 2$. Non ci sono ricette infallibili, ma è importante acquisire esperienza ed esercitare l'intuizione fisica - l'arma che di solito manca all'analista puro. Tra gli strumenti moderni più potenti consideriamo qui 1) l'utilizzazione di trasformate integrali - tipicamente la trasformata di Fourier, 2) la tecnica di *splitting* di cui la formula di Lie-Trotter è un esempio e 3) le tecniche di *matrici sparse* che permettono di affrontare problemi di grande taglia ($N \simeq O(10^{12})$).

4.5.1. Implementazione numerica della formula di Lie-Trotter. Sia $\mathcal{H} = \mathcal{K} + \mathcal{V}$ l'Hamiltoniano di un sistema quantistico a un grado di libertà, $\mathcal{K} = \frac{1}{2}p^2$ e $\mathcal{V} = V(q)$. Lo stato quantistico $|t=0\rangle$ sarà rappresentato da una “funzione d'onda” $\psi(x, 0) = \langle x|t=0\rangle$. Si tratta di costruire la soluzione dell'equazione $i \frac{d}{dt} \psi(x, t) = \mathcal{H} \psi(x, t)$ noto lo stato iniziale $\psi(x, 0)$. Partiamo dalla formula (4.12) che suggerisce un metodo per il calcolo della soluzione: fissiamo un intervallo temporale massimo su cui studiare il problema $0 \leq t \leq T$ e scegliamo $N = N_t$ nella formula di Trotter “sufficientemente grande”: qui la condizione da rispettare è $ET/N_t \ll 1$ se si vuole trattare accuratamente funzioni d'onda con energia $\sim E$. La funzione $\psi(x)$ dovrà poi essere “campionata” su un insieme finito di punti, la scelta più semplice essendo quella di fissare un reticolo di punti $x_i = -L/2 + i\alpha$ con $\alpha = L/N_x$. N_x dovrà essere fissato abbastanza grande da rendere accettabile l'errore “ultravioletto”, che consiste nel fatto che solo i momenti $p \leq \pi/\alpha$ vengono riprodotti accuratamente. Il bilanciamento tra tutti questi parametri determina l'accuratezza dell'algoritmo.

Il calcolo si imposta poi in modo iterativo; all' n -esimo passo si procede secondo lo schema

- i) $\psi_{\text{tmp}}(x_i) = e^{-i\tau V(x_i)} \psi(x_i, n\tau)$, ($i = 1, \dots, N_x$)
- ii) $\psi(x, (n+1)\tau) = e^{-i\tau \mathcal{K}} \psi_{\text{tmp}}$

Il secondo passo è problematico, in quanto l'applicazione della formula (4.13) comporta un doppio loop e cioè un numero di operazioni che cresce come N_x^2 . Nel caso di problemi tridimensionali

questo fatto può rendere del tutto impraticabile l'idea - immaginiamo $N_x \sim O(10^3)$, ciò implicherebbe $O(10^{18})$ operazioni aritmetiche - del tutto improponibile. La soluzione, nota da tempo [FFS82, OMT91] consiste nel modificare l'iterazione inserendo due trasformate di Fourier, schematicamente

$$\exp\{-it\mathcal{K}/N\}\exp\{-it\mathcal{V}/N\} \hookrightarrow \mathcal{F}^\dagger \exp\{-it\mathcal{K}(p)/N\}\mathcal{F} \exp\{-it\mathcal{V}(x)/N\}$$

Il risultato è il nuovo schema iterativo

- i) $\psi_{\text{tmp}}(x_i) \leftarrow e^{-it\mathcal{V}(x_i)} \psi(x_i, n\tau)$, ($i = 1, \dots, N_x$)
- ii) $\phi(p_j) \leftarrow (\mathcal{F} \psi_{\text{tmp}})(p_j)$
- iii) $\phi(p_j) \leftarrow e^{-it\mathcal{K}(p_j)} \phi(p_j)$
- iv) $\psi(x, (n+1)\tau) \leftarrow \mathcal{F}^\dagger \phi$

Naturalmente anche la trasformata di Fourier, se discretizzata in modo ingenuo, comporta $O(N_x^2)$ operazioni aritmetiche; esiste invece la possibilità di calcolare la trasformata in $O(N_x \log(N_x))$ operazioni, adottando la *trasformata di Fourier veloce* o FFT. In questo modo risulta realizzabile l'idea della formula di Trotter anche in sistemi a 3 gradi di libertà. I programmi `matlab` in allegato `wms`, `wms2`, `wavepktj` realizzano l'idea in 1 e 2 gradi di libertà. In figura il caso di un potenziale $V(x) \propto (x^2 - x_0^2)^2$ con una funzione iniziale concentrata intorno al “vuoto classico” di sinistra.

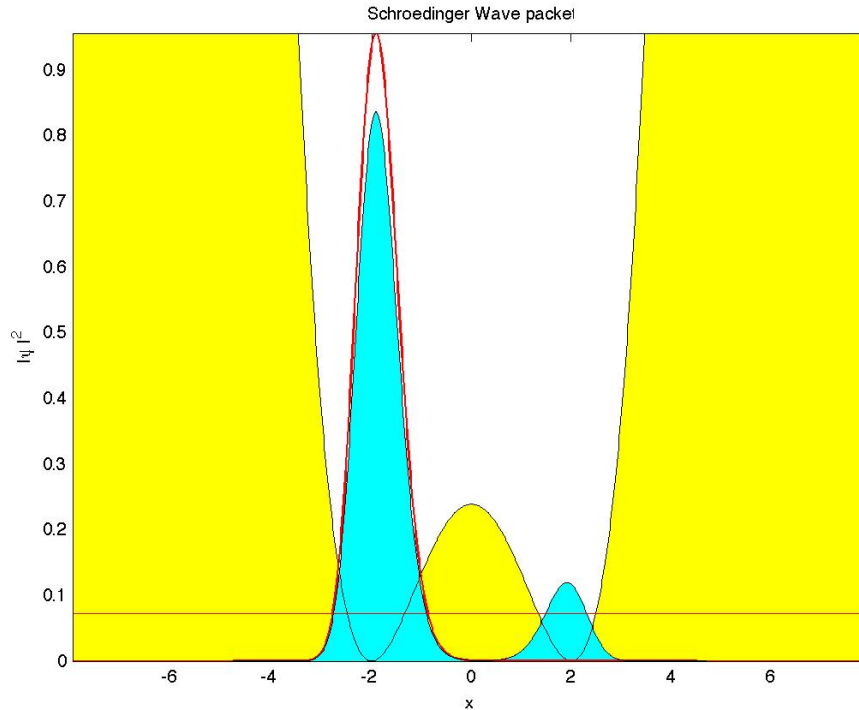


FIGURA 4-4. La formula di Lie-Trotter al lavoro.

Il nucleo del programma è dato da

```

while(t < t_max),
    psi = psi.*Uhalf;
    phi = fft(psi); % to momentum space
    phi = phi.*W;
    psi = ifft(phi); % back
    psi = psi.*Uhalf;
    t = t+tau;
    ...
end

```

Qui W contiene $\exp\{-i\tau p^2/2\}$ ed è calcolato una sola volta prima di iniziare l'iterazione; $U_{1/2}$ contiene l'evoluzione in x ma di un tempo $\tau/2$, e infatti viene applicato due volte: la scelta simmetrizzata $\exp\{-i\tau V/2\} \exp\{-i\tau K\} \exp\{-i\tau V/2\}$ è il più semplice sistema di “miglioramento” della formula di Lie–Trotter, con un errore complessivo $O(\tau^2)$ anziché $O(\tau)$.^{*} Questo accorgimento appesantisce il calcolo ed è adottato se nel corso del loop si opera sulla funzione ψ (plot, misure varie, etc). Altrimenti è sufficiente adottare uno schema più economico, raccogliendo il “miglioramento” alla fine:

```

psi = psi.*conj(Uhalf);
while(t < t_max),
    phi = fft(psi.*U);
    psi = ifft(phi.*W);
    t = t+tau;
end
psi = psi.*Uhalf;

```

Si può consultare a Pag. 146 il codice `matlab` che prepara i vettori x_i e p_i .

🔗 La disponibilità di trasformate veloci, oltre alla FFT, allarga il dominio di applicabilità della formula di Lie–Trotter. È possibile affrontare problemi di evoluzione temporale per funzioni definite sulla sfera $S^2 = \{x \in \mathbb{R}^3 \mid \|x\| = 1\}$ utilizzando una versione veloce della trasformata[†] $\psi(\vartheta, \varphi) \approx c_{\ell m}$

$$\psi(\vartheta, \varphi) = \sum_{\substack{\ell \geq 0 \\ -\ell \leq m \leq \ell}} c_{\ell m} \mathcal{Y}_{\ell}^m(\vartheta, \varphi)$$

Utilizzando questa trasformata veloce è possibile affrontare efficientemente problemi di evoluzione sulla sfera con Hamiltoniano $H = -\Delta + V(x)$ dove Δ è l'operatore di Laplace–Beltrami sulla sfera (codice `matlab` nella directory `mexs2kit`). ◀

4.5.2. Matrici sparse e problemi spettrali. Per una trattazione completa dell'approccio numerico a problemi spettrali, o più in generale alla soluzione di problemi lineari, il riferimento principale è tuttora [GL96]. Qui mi preme illustrare un aspetto particolarmente importante e di cui tutti possono trarre vantaggio anche senza dovere trasformarsi in esperti analisti. L'elemento cruciale che permette di affrontare problemi di grande taglia è rappresentato dalla introduzione della struttura di “matrice sparsa”. L'idea consiste nello sfruttamento di una caratteristica tipica delle matrici che entrano nei problemi di fisica matematica una volta discretizzati e che discende dal carattere locale delle interazioni: le matrici che rappresentano l'energia ovvero le forze di un sistema dinamico sono in genere a bassa densità di elementi di matrice, ossia una porzione rilevante di elementi di matrice sono nulli. Ad es. la forza di una catena di oscillatori accoppiati,

^{*}L'implementazione della FFT adottata in `matlab` è al momento la *FFTW* di Frigo, Johnson e Kral., ([fftw@fftw.org](http://fftw.org))

[†]Parole chiave: SpharmonicKit, S2kit.

Eq. 1.4 è diversa da zero solo lungo la diagonale principale e le due diagonali immediatamente adiacenti - ossia si tratta di una matrice *tridiagonale*; in generale, in dimensionalità più alta, le matrici interessanti non sono tridiagonali ma pure presentano una percentuale di elementi di matrici non-nulli piccola rispetto al totale. tipicamente $O(dN^d)$ rispetto a $O(N^{2d})$, se d è la dimensione dello spazio in cui agiscono le interazioni. Una matrice a bassa densità è detta in gergo *sparsa* e per questi casi si è introdotto un metodo di rappresentazione che sfrutta queste caratteristiche per risparmiare spazio in memoria e velocizzare le operazioni aritmetiche. La struttura è del tipo

riga	colonna	elemento di matrice
i_1	j_1	M_{i_1,j_1}
i_2	j_2	M_{i_2,j_2}
i_3	j_3	M_{i_3,j_3}
		\dots
i_r	j_r	M_{i_r,j_r}

e si intende che tutti gli elementi che non figurano nella lista sono uguali a zero. Esistono routines aritmetiche che realizzano le operazioni basilari di somma, prodotto, trasposta, etc. I linguaggi `matlab` e `mathematica` implementano questa struttura e permettono di costruire facilmente matrici sparse e condurre tutte le operazioni su di esse (vedi ad es. il Probl. 25).

Per quanto riguarda lo studio dello spettro, l'utilizzazione di matrici sparse permette di considerare matrici di grandi dimensioni ed ottenere così un calcolo più accurato: la dimensione finita della matrice infatti rappresenta un elemento di approssimazione in problemi di fisica del continuo che richiederebbero la soluzione di problemi spettrali nello spazio a infinite dimensioni; tanto più grande è la dimensione della matrice tanto più ci si avvicina al problema reale. La routine `eigs` di `matlab`, basata sull'algoritmo di Arnoldi, permette di ricavare autovalori ed autovettori relativi a parte dello spettro - tipicamente la parte di bassa energia nel caso di meccanica quantistica.

Un aspetto interessante dell'algoritmo impiegato, tutt'altro che banale, consiste nel fatto che l'unica informazione richiesta per il funzionamento dell'algoritmo è rappresentata dalla azione dell'operatore su un generico vettore; si richiede cioè la regola per costruire $\mathcal{M}\xi$ per un arbitrario vettore ξ . Questa regola può essere definita dalla assegnazione di \mathcal{M} come matrice sparsa; ma, e qui sta il fatto che rende tutto lo schema così potente, si può in alternativa chiamare una routine esterna che ritorni $\mathcal{M}\xi$ in output. Questa routine ha un formato del tutto arbitrario e può contenere operazioni complesse quali trasformate integrali, interpolazioni etc. La routine `eigs` ammette chiamate in cui il primo argomento è una matrice oppure una routine. Si vedano gli esercizi per una descrizione dettagliata di qualche esempio. Una tipica applicazione di questa strategia di calcolo consiste nella rappresentazione di operatori differenziali attraverso la trasformata di Fourier

$$p^2 \psi = -\frac{d^2 \psi}{dx^2} \equiv \mathcal{F}^\dagger (p^2 \mathcal{F}(\psi))$$

il che permette una rappresentazione molto accurata della derivata e al tempo stesso, attraverso la routine FFT, una notevole efficienza del codice.

4.5.3. L'algoritmo di Lanczos. Il prototipo di routine "sparsa" per il calcolo dello spettro risale ai primi anni '50 ed è dovuta a C. Lanczos [Lan50]. L'idea è quella di sostituire al problema di determinare una trasformazione di similitudine

$$\mathcal{M} = S \Lambda S^{-1}$$

con Λ diagonale quello, più limitato, di determinare una trasformazione

$$\mathcal{M} = S \mathcal{T} S^{-1}$$

tale che \mathcal{T} sia *tridiagonale*. Una volta costruita la \mathcal{T} il suo spettro si può determinare in modo molto efficiente con il metodo del criterio di Sturm[‡]. Vediamo: sia

$$\mathcal{T} \leftrightarrow \begin{pmatrix} \alpha_1 & \beta_1 & 0 & \dots \\ \beta_1 & \alpha_2 & \beta_2 & 0 & \dots \\ 0 & \beta_2 & \alpha_3 & \beta_3 & 0 & \dots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \end{pmatrix}$$

inoltre identifichiamo le colonne di \mathcal{S} con n vettori $\|e^{(1)}, e^{(2)}, \dots, e^{(n)}\|$, ossia $S_{ij} \equiv e_i^{(j)}$. Nel caso simmetrico, i vettori $e^{(j)}$ sono tra loro ortogonali. Allora la relazione $\mathcal{S}\mathcal{T} = \mathcal{M}\mathcal{S}$ diventa

$$\mathcal{M}e^{(j)} = \beta_{j-1}e^{(j-1)} + \alpha_j e^{(j)} + \beta_j e^{(j+1)}$$

Da questa relazione si ottiene $e^{(j+1)}$ per ricorrenza a partire da $e^{(1)}$ che può essere scelto arbitrariamente. Tenendo conto di ortogonalità e normalizzazione si ricava successivamente

$$\mathcal{M}e^{(1)} = \alpha_1 e^{(1)} + \beta_1 e^{(2)} \Rightarrow \alpha_1 = \langle e^{(1)}, \mathcal{M}e^{(1)} \rangle$$

$$e^{(2)} = \beta_1^{-1} (\mathcal{M} - \alpha_1 \mathbb{1}) e^{(1)}$$

(β_1 è fissato dalla condizione $\|e^{(2)}\| = 1$)

$$\mathcal{M}e^{(2)} = \beta_1 e^{(1)} + \alpha_2 e^{(2)} + \beta_2 e^{(3)} \Rightarrow \alpha_2 = \langle e^{(2)}, \mathcal{M}e^{(2)} \rangle$$

$$e^{(3)} = \beta_2^{-1} [(\mathcal{M} - \alpha_2 \mathbb{1}) e^{(2)} - \beta_1 e^{(1)}]$$

...

Purtroppo l'algoritmo non è utilizzabile in pratica per via del fatto che l'ortogonalità dei vettori $e^{(j)}$ è verificata solo in aritmetica esatta; lavorando in precisione finita, invece, si genera un errore che rapidamente si amplifica fino a degradare completamente la natura ortonormale della base. Ciò ha effetti imprevedibili sullo spettro, ma con grande probabilità assolutamente letali (per esempio si producono copie spurie di autovalori multipli). Varianti dell'algoritmo vengono incontro a questo problema e sono correntemente utilizzate per calcoli realistici; la routine `eigs` di `matlab` impiega l'algoritmo di Arnoldi, della libreria `Arpack` (si veda <http://www.caam.rice.edu/software/ARPACK/> dove si trova ampia documentazione). Arnoldi ha il vantaggio di un'ottima stabilità e può essere impiegato anche a operatori non-simmetrici. Se si vuole implementare il proprio codice facendo uso solo di software "open source" la disponibilità di `Arpack`, `FFTW`, `gsl`, `etc.` è da tenere ben presente, in un quadro in continua evoluzione. Per un "case study" che illustra bene un'applicazione di `Arpack` e di varie trasformate integrali si possono consultare i documenti <http://arxiv.org/abs/physics/0407098>, <http://arxiv.org/abs/math-ph/0407021>.

ESEMPIO 4-20. Diamo un esempio di calcolo di spettro in meccanica quantistica pienamente documentato. Altri esempi sono da cercare nella sezione Esercizi. Si tratta di calcolare lo spettro di energia per una particella scalare immersa in un campo di forze con energia potenziale $V(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$. L'operatore Hamiltoniano è (in unità opportune)

$$\mathcal{H} = -\frac{1}{2} \Delta + V(\mathbf{x})$$

e si intende rappresentare il Laplaciano Δ utilizzando la trasformata di Fourier. I passi per costruire il programma sono i seguenti:

- (a) scegliere alcuni parametri "tecnici", numero di autovalori desiderato, lato della scatola di normalizzazione ($|x| < L$, $|y| < L$, $|z| < L$) eventualmente utilizzando un parallelepipedo con lati differenti nel caso sia conveniente per la particolare scelta di $V(\mathbf{x})$, il numero di punti del reticolo discreto che si introduce per campionare la funzione d'onda;

[‡]Si veda [GL96]. Un esempio è fornito nel codice "tridiag.c".

- (b) costruire il reticolo in \mathbf{x} e nello spazio dei momenti:

```

Reticolo x e p
dx = 2*L/N;
x0 = -L*(1-1/N) : dx : L*(1-1/N);
[x,y,z] = meshgrid(x0,x0,x0);
r = sqrt(x.^2+y.^2+z.^2);

k = fftshift(pi*(-N/2:N/2-1)/L);
[kx,ky,kz]=meshgrid(k,k,k);
k2 = kx.^2 + ky.^2 + kz.^2;

```

- (c) definire operatori differenziali di utilizzo comune (nello spazio dei momenti): $\partial^2/\partial x^2 \leftrightarrow -k_x.^2$, etc.; $\Delta \leftrightarrow -k_x.^2 - k_y.^2 - k_z.^2$.
- (d) definire l'energia potenziale, ad es $V = r.^2/2$;
- (e) definire l'operatore Hamiltoniano come sotto-routine: il vettore di ingresso è sempre per **eigs** un vettore colonna; essendo “psi” un array $N \times N \times N$ è necessario trasformarlo attraverso l'istruzione “reshape”; “fftn” riconosce l'array 3-D e ne prende la trasformata di Fourier; la moltiplicazione per $-k^2$ realizza il Laplaciano nello spazio dei momenti e la successiva antitrasformata ci riporta nello spazio delle coordinate; un ultimo “reshape” è necessario per inviare a **eigs** un vettore colonna:

```

Hamiltoniana
function xout=Hamiltonian(xin)
    psi = reshape(xin(1:N^3),N,N,N);
    fpsi = fftn(psi);
    Deltapsi = real(ifftn(k2.*fpsi));
    Hpsi = 0.5* Deltapsi + V .* psi;
    xout = reshape(Hpsi, N^3, 1);
end

```

Notare che la routine che definisce l'Hamiltoniana è preferibilmente “nested” di modo che eredita tutti i parametri dalla funzione principale - anche se ciò richiede molta attenzione per evitare conflitti accidentali nella denominazione delle variabili;


- (f) chiamare **eigs** con le opportune opzioni (“SA” significa “smallest algebraic”, dunque gli autovalori vicini allo stato fondamentale):

```

schr3D.m (frammento)
opts.tol = 1e-10; % accuracy
opts.disp = 0; % quiet!
opts.isreal = true;
opts.issym = true;
[Psi, E] = eigs(@Hamiltonian, N^3, Neig, 'SA', opts);

```

Problemi

 Il simboli * e ** indicano gli esercizi più impegnativi. Il simbolo # segnala che la soluzione, almeno parziale, è riportata in fondo al capitolo.

PROB 1. *Dimostrare che ℓ_2 è uno spazio lineare.*

PROB2. *Dimostrare che $C_2(\mathbf{a}, \mathbf{b})$ e $H_2(\mathbb{D})$ sono spazi lineari (vedi a Pag. 107 per le definizioni).*

PROB 3. Per quali valori di α la successione $\{n^{-\alpha} | n = 1, 2, 3, \dots, \infty\}$ è in ℓ_2 ?

PROB 4. *Determinare quale tra le seguenti funzioni individua un vettore di $L_2(0,1)$:*

- (a) $(1 + x^2)^{-1/2}$
 (b) $1/\cos(\omega x)$
 (c) $x^\alpha, \quad \alpha \in \mathbb{R}$

PROB 5.^{#*} Si consideri la successioni di vettori in ℓ_2 :

$$\begin{aligned} e_1 &= (a, b, 0, 0, 0, 0, \dots, 0, \dots) \\ e_2 &= (0, a, b, 0, 0, 0, \dots, 0, \dots) \\ e_3 &= (0, 0, a, b, 0, 0, \dots, 0, \dots) \\ &\vdots \\ e_n &= (\underbrace{0, 0, \dots, 0}_{n-1}, a, b, 0, 0, \dots, 0, \dots) \\ &\vdots \end{aligned}$$

Si chiede: per quali valori di (a, b) i vettori e_i formano una base in ℓ_2 ?

PROB 6.[#] Costruire un esempio di funzione continua appartenente a $L_2(-\infty, \infty)$ e che non tenda a zero per $x \rightarrow \pm\infty$. Dimostrare poi che se $f(x)$ è assolutamente continua, con $f'(x) \in L_2(-\infty, \infty)$ allora $f(x) \rightarrow 0$ per $x \rightarrow \pm\infty$.

PROB 7.^{#*} L'operatore lineare \mathfrak{G} è definito su una base ortonormale $|\mathfrak{n}\rangle$ dalle relazioni

$$\begin{aligned} \mathfrak{G}|1\rangle &= |1\rangle \\ \mathfrak{G}|2\rangle &= |1\rangle, \quad \mathfrak{G}|3\rangle = |2\rangle \\ \mathfrak{G}|4\rangle &= |1\rangle, \quad \mathfrak{G}|5\rangle = |2\rangle, \quad \mathfrak{G}|6\rangle = |3\rangle \\ \mathfrak{G}|7\rangle &= |1\rangle, \quad \mathfrak{G}|8\rangle = |2\rangle, \quad \mathfrak{G}|9\rangle = |3\rangle, \quad \mathfrak{G}|10\rangle = |4\rangle \\ &\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ \mathfrak{G}|k_n\rangle &= |1\rangle, \quad \mathfrak{G}|k_n+1\rangle = |2\rangle, \quad \dots\dots\dots, \quad \mathfrak{G}|k_n+n-1\rangle = |n\rangle \\ &\vdots \end{aligned}$$

$k_n = \frac{1}{2}(n^2 - n + 2)$, $n = 1, 2, 3, \dots$. Dimostrare che \mathfrak{G} è illimitato e che il dominio di \mathfrak{G}^\dagger è vuoto (vedere [Hal82] #50).

PROB 8. *Studiare l'operatore differenziale \mathcal{X} definito da*

$$(\mathcal{X}f)(x) = -\frac{1}{2} \left(x \frac{d}{dx} + \frac{d}{dx} x \right) f(x)$$

con dominio $D_{\mathcal{X}} = \{f \in L_2(0, \infty) \mid x f'(x) \in L_2\}$. Determinare se \mathcal{X} è autoaggiunto e calcolare $(\exp\{it\mathcal{X}\}f)(x)$.

PROB 9.[#] *Sia \mathcal{Y} l'operatore in $L_2(a, b)$ con $0 < a < b$,*

$$(\mathcal{Y}f)(x) = \frac{1}{2}i \left(x^2 \frac{df(x)}{dx} + \frac{d}{dx} (x^2 f(x)) \right)$$

definito per ogni f a.c. con $f' \in L_2(a, b)$. e con le condizioni al contorno

$$a f(a) = b f(b)$$

Si determini lo spettro di \mathcal{Y} .

PROB 10.[#] *Sia \mathcal{D} l'operatore definito sulle funzioni in $\mathfrak{H} = L_2(-\infty, \infty; \exp(-x^2) dx)$ che siano a.c., con derivata in \mathfrak{H} come $\mathcal{D}f(x) = df(x)/dx$. Si determini l'aggiunto Hermitiano di \mathcal{D} , lo spettro di \mathcal{D} e quello di $\mathcal{D}^\dagger \mathcal{D}$.*

PROB 11.^{#*} *Siano p e q gli operatori canonici del Cap.4.3.3. Dimostrare l'identità*

$$\exp\{-t(p^2 + iq)\} = e^{-itq/2} e^{-tp^2} e^{-itq/2} e^{-t^3/12}$$

e dedurre che l'operatore $\mathcal{A} = p^2 + iq$ ha spettro vuoto.

PROB 12.^{#*} *Sia \mathcal{V} (in omaggio a Vito Volterra) l'operatore in $L_2(0, 1)$*

$$(\mathcal{V}f)(x) = \frac{1}{\sqrt{\pi}} \int_0^x \frac{f(y)}{\sqrt{x-y}} dy$$

- (a) *Dimostrare che \mathcal{V} è limitato;*
- (b) *determinare \mathcal{V}^{-1}*
- (c) *determinare lo spettro di \mathcal{V} .*

PROB 13.[#] *Sia \mathcal{W}_α l'operatore in $L_2(0, 1)$*

$$(\mathcal{W}_\alpha f)(x) = \frac{1}{\sqrt{\pi}} \int_0^x \frac{f(y)}{(x-y)^\alpha} dy, \quad (0 < \alpha < 1)$$

- (a) *Dimostrare che \mathcal{W}_α è limitato;*
- (b) *determinare \mathcal{W}_α^{-1}*

PROB 14.[#] *Sia $(p^2 f)(x) = -f''(x)$ con c.c. periodiche ($f(-\pi) = f(\pi), f'(-\pi) = f'(\pi)$). Calcolare il risolvete di p^2 .*

PROB 15. *Come il problema precedente con differenti condizioni al contorno:*

- i) $f(-\pi) = f(\pi) = 0$
- ii) $f(-\pi) = -f(\pi), \quad f'(-\pi) = -f'(\pi)$

PROB 16.[#] Siano \mathbf{q} e \mathbf{p} gli operatori canonici in $L_2(-\infty, \infty)$. Se indichiamo con $|\mathcal{A}|$ il modulo di \mathcal{A} , cioè la funzione di \mathcal{A} che corrisponde al valore assoluto $x \rightarrow |x|$, determinare la matrice rappresentativa dell'operatore $\mathcal{S} = |\mathbf{p}| + |\mathbf{q}|$ nella base dei polinomi di Hermite $u_n(x) = (\sqrt{\pi} 2^n n!)^{-1/2} H_n(x) e^{-x^2/2}$. Dimostrare che in base a semplici proprietà di simmetria

$$\begin{aligned}\langle u_{m+1}, \mathcal{S} u_{2n} \rangle &= 0 \\ \langle u_{4m+1}, \mathcal{S} u_{4n+3} \rangle &= 0 \\ \langle u_{4m}, \mathcal{S} u_{4n+2} \rangle &= 0\end{aligned}$$

Studiare lo spettro di \mathcal{S} .

PROB 17. Si definisce come raggio spettrale di un operatore limitato \mathcal{A} il limite superiore

$$r(\mathcal{A}) = \sup\{|z| \mid z \in \sigma(\mathcal{A})\}$$

Dimostrare che $r(\mathcal{A}) \leq \sup_n \|\mathcal{A}^n\|^{1/n}$.

PROB 18.[#] Dimostrare che per due operatori limitati \mathcal{A} e \mathcal{B} se $\mathcal{C} = [\mathcal{A}, \mathcal{B}]$ commuta con entrambi gli operatori allora lo spettro di \mathcal{C} è costituito al più del punto $z = 0$.

PROB 19. Sia $H_2(\mathbb{D})$ lo spazio di Hilbert delle funzioni analitiche regolari $f(z)$ nel disco $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$ tali che

$$\iint_{\mathbb{D}} |f(x + iy)|^2 dx dy < \infty$$

Per ogni funzione analitica $h(z)$ meromorfa all'interno del disco si definisce $(\mathcal{T}_h f)(z) \equiv h(z) f(z)$. Discutere le principali proprietà (definizione del dominio, invertibilità, etc) di \mathcal{T}_h in particolare per $h(z) = z^n$, $n \in \mathbb{Z}$.

PROB 20.[#] Dimostrare che il funzionale lineare $\delta_z(f) = f(z)$ in $H_2(\mathbb{D})$ è continuo e determinare il vettore $e_z \in H_2(\mathbb{D})$ tale che

$$f(z) = \langle e_z, f \rangle = \iint_{\mathbb{D}} \overline{e_z(w)} f(w) dw \wedge d\bar{w}$$

PROB 21. Lo spazio di Bargmann \mathfrak{B} è lo spazio di Hilbert delle funzioni $f(z)$ analitiche regolari in tutto \mathbb{C} tali che esista l'integrale

$$\frac{1}{\pi} \iint_{\mathbb{C}} \exp\{-|z|^2\} |f(z)|^2 dz \wedge d\bar{z}$$

- (a) Dimostrare che il funzionale lineare $\delta_z(f) = f(z)$ è continuo;
- (b) determinare l'aggiunto Hermitiano dell'operatore $(\mathbf{a} f)(z) = \frac{df}{dz}$;
- (c) determinare lo spettro dell'operatore

$$\mathcal{H} = \mathbf{a}^\dagger \mathbf{a} + c \mathbf{a} + \bar{c} \mathbf{a}^\dagger$$

con c costante arbitraria;

- (d)^{#*} calcolare la soluzione dell'equazione differenziale

$$i \frac{d\psi(t)}{dt} = (\mathbf{a}^\dagger \mathbf{a} + c \mathbf{a} e^{i\omega t} + \bar{c} \mathbf{a}^\dagger e^{-i\omega t}) \psi(t)$$

con $\psi \in \mathfrak{B}$.

(e)^{#*} calcolare la soluzione dell'equazione differenziale

$$i \frac{d\psi(t)}{dt} = (a^\dagger a + c a^2 e^{2i\omega t} + \bar{c} a^{\dagger 2} e^{-2i\omega t}) \psi(t)$$

con $\psi \in \mathfrak{B}$. In che cosa si differenzia qualitativamente la soluzione di questo problema dal precedente?

(f) L'operatore $\mathcal{V}_t = e^{t d/dz}$ è illimitato, come si verifica facilmente applicandolo alle funzioni $e^{\lambda z}$. D'altra parte $(\mathcal{V}_t f)(z) = f(z+t)$ per ogni funzione intera f e per ogni t . Esiste $f(z) \in \mathfrak{B}$ tale che $f(z+t) \notin \mathfrak{B}$? Dimostrare che modificando \mathcal{V} come segue

$$(\mathcal{U}_t f)(z) = e^{-1/2 t^2 - \bar{t} z} f(z+t)$$

si ottiene un gruppo di trasformazioni unitarie in \mathfrak{B} .

PROB 22.[#] Siano p e q gli operatori canonici rappresentati in $L_2(-\infty, \infty)$ da

$$(q\psi)(x) = x\psi(x); \quad (p\psi)(x) = -i \frac{d\psi(x)}{dx}$$

Determinare lo spettro di $\mathcal{A} = 2^{-1/2}(q + ip)$; si trovi la trasformazione $\mathcal{K} : L_2 \mapsto \mathfrak{B}$ (\mathfrak{B} è lo spazio di Bargmann) tale che $\mathcal{K}\mathcal{A} = a\mathcal{K}$ (dove a è definito al punto (b) del problema precedente).

PROB 23.[#] Sia $|n\rangle$ una base ortonormale nello spazio di Hilbert. Si definisce l'operatore \mathcal{A}_q per ogni reale q , ($0 < q < 1$) come segue

$$\mathcal{A}_q |n\rangle = \begin{cases} 0 & (n=0) \\ \sqrt{1-q^{2n}} |n-1\rangle & (n>0) \end{cases}$$

Dimostrare le seguenti proprietà di \mathcal{A}_q :

- (a) \mathcal{A}_q è limitato;
- (b) $[\mathcal{A}_q, \mathcal{A}_q^\dagger] = (1-q^2)(\mathbb{1} - \mathcal{A}_q^\dagger \mathcal{A}_q)$;
- (c)

$$\mathcal{A}_q^n \mathcal{A}_q^{\dagger m} = \sum_{j \geq 0} [j]! \begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} m \\ j \end{bmatrix} q^{2(n-j)(m-j)} \mathcal{A}_q^{\dagger m-j} \mathcal{A}_q^{n-j}$$

- (d) \mathcal{A}_q ha spettro puntuale $\{\lambda \in \mathbb{C} \mid |\lambda| < 1\}$, con autovettori $|\lambda\rangle = \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{[n]!}} |n\rangle$. (I simboli $[n]!$ e $\begin{bmatrix} n \\ j \end{bmatrix}$ sono stati definiti nell'eserc. II.33).

PROB 24.^{#*} Sia H_2^ℓ lo spazio di Hilbert delle funzioni analitiche regolari nel disco $\{|z| < 1\}$ con prodotto interno

$$\langle f, g \rangle \equiv \frac{2\ell-1}{\pi} \iint_{|z|<1} (1-|z|^2)^{2\ell-2} \overline{f(z)} g(z) dz \wedge d\bar{z}, \quad (\ell > 1/2).$$

- i) Dimostrare che $\delta_z(f) = f(z)$ è un funzionale lineare continuo.
- ii) Se $(\mathcal{Q}f)(z) = z f(z)$ determinare il suo aggiunto \mathcal{Q}^\dagger .
- iii) Costruire l'operatore $\exp\{t d/dz\}$ e discuterne il dominio.
- iv) Dimostrare che

$$(\mathcal{U}f)(z) = (\bar{b}z + \bar{a})^{-s} f\left(\frac{az+b}{\bar{b}z+\bar{a}}\right)$$

con $|a|^2 - |b|^2 = 1$ definisce un operatore unitario per un'opportuna scelta del parametro s .

PROB 25.^{#**} Si studi lo spettro della matrice \mathcal{M} definita da

$$\begin{aligned}\langle 0, n | \mathcal{M} | 0, n \rangle &= (1 + \lambda(1 - \nu \delta_{n1})) n \\ \langle 0, n+1 | \mathcal{M} | 0, n \rangle &= \sqrt{\lambda n(n+1)}\end{aligned}$$

nei due casi $\nu = 0$ e $\nu = 1$.

PROB 26. Sia \mathcal{J} l'operatore in $H_2(\mathbb{D})$ (vedi Probl. 19) definito da

$$(\mathcal{J}f)(z) = z \frac{df}{dz}.$$

La relazione di commutazione $[\frac{d}{dz}, z \frac{d}{dz}] = \frac{d}{dz}$ implica che

$$\mathcal{J}_t = e^{t \frac{d}{dz}} \mathcal{J} e^{-t \frac{d}{dz}} = \mathcal{J} + t \frac{d}{dz}$$

Si chiede: è lecito concludere che $\mathcal{J}_0 = \mathcal{J}$ e \mathcal{J}_t hanno lo stesso spettro? Dopo avere tratto le conclusioni ... calcolare lo spettro di \mathcal{J}_t .

PROB 27. Sia \mathbf{a} l'operatore di "annichilazione" definito su una base ortonormale da $\mathbf{a}|n\rangle = \sqrt{n}|n-1\rangle$. Si determini lo spettro dell'operatore

$$\mathcal{S} = (1 + \mathbf{a}^\dagger \mathbf{a})^{-1/2} \mathbf{a}$$

e si mostri che \mathcal{S}^\dagger è isometrico ma non unitario.

PROB 28.[#] L'operatore \mathfrak{G} è definito in $L_2(-\infty, \infty)$ dall'espressione

$$(\mathfrak{G}f)(x) = \frac{1}{\pi} \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} \frac{f(x-y) - f(x)}{y^2 + \varepsilon^2} dy$$

con $f(x)$ a.c., $f'(x) \in L_2$. Determinare $\exp t\mathfrak{G}$.

PROB 29.[#] Sia $f \in L_2(-1, 1; w(x)dx)$ sufficientemente regolare affinché esista l'integrale (in parte principale di Cauchy)

$$(\mathcal{K}f)(x) = \int_{-1}^1 \frac{f(x) - f(y)}{(x-y)^2} w(y) dy$$

Determinare lo spettro puntuale di \mathcal{K} nel caso $w(x) = \sqrt{1-x^2}$.

PROB 30.^{#*} Si studi l'operatore integrale (l'operatore di **Tuck**^{*}, vedi [TN02])

$$(4.14) \quad (\mathcal{K}\psi)(\xi) = \int_0^1 d\eta \frac{\psi(\eta) - \psi(\xi)}{|\xi - \eta|}$$

e si mostri che \mathcal{K} è diagonalizzabile nella base dei polinomi di Legendre $P_n(2\xi-1)$ con autovalori dati dai numeri armonici $h_n = \sum_{j=1}^n j^{-1}$.

^{*}Ernest O. Tuck, illustre matematico australiano, Adelaide 1939–2009.

PROB 31.^{#*} Si studi l'operatore integrale

$$(\mathcal{Z}f)(x) = \frac{1}{\pi} \frac{d}{dx} \int_{-1}^1 \sqrt{1-y^2} \frac{f(y)}{y-x} dy$$

quale operatore in un opportuno dominio $\in L_2((-1, 1), \sqrt{1-x^2})$. Determinare lo spettro di \mathcal{Z} (suggerimento: considerare per prima cosa una $f(x)$ polinomiale).

PROB 32. Si consideri l'operatore lineare

$$(\Delta f)(x) = -\frac{1}{2}h^{-2} (f(x+h) - 2f(x) + f(x-h))$$

dove $f \in L_2(-\pi, \pi)$, $h > 0$ e $x+h$ o $x-h$ sono da considerare modulo 2π . Dimostrare che Δ è limitato e che il suo spettro è dato da

$$\sigma(\Delta) = \left\{ \frac{1 - \cos(nh)}{h^2} \mid n \in \mathbb{Z} \right\}$$

PROB 33. Sia \mathcal{D} l'operatore $-i \frac{d}{dx}$ definito in $L_2(-\pi, \pi)$ con condizioni periodiche al contorno. Si determini il nucleo integrale $K(x, x'; t)$ tale che

$$(\exp\{-t|\mathcal{D}|\} f)(x) = \int_{-\pi}^{\pi} K(x, x'; t) f(x') dx'$$

intendendo al solito $|\text{Op}|$ il valore assoluto di Op . Derivando rispetto a t si ricavi la forma esplicita di $|\mathcal{D}|$.

PROB 34.[#] Sia $\psi(\theta)$ una qualunque funzione a.c. con $\psi' \in L_2$ e tale che $\psi(\pi) = \psi(-\pi) + 2n\pi$ per n intero. Si trovi il minimo del funzionale (non-lineare)

$$\mathcal{H}[\psi] \equiv \frac{1}{4\pi^2} \iint_{-\pi}^{\pi} \left| \frac{e^{i\psi(\theta)} - e^{i\psi(\theta')}}{e^{i\theta} - e^{i\theta'}} \right|^2 d\theta d\theta'$$

PROB 35.[#] $n \neq \infty$ Sia \mathcal{T} l'operatore "di shift" in ℓ_2

$$\mathcal{T} |n\rangle = \begin{cases} 0 & (n=1) \\ |n-1\rangle & (n>1) \end{cases}$$

Determinare lo spettro dell'operatore autoaggiunto $\mathcal{X} = \frac{1}{2}(\mathcal{T} + \mathcal{T}^\dagger)$. Sia \mathcal{T}_N il troncamento di \mathcal{T} al sottospazio finito generato da $|n\rangle, n \leq N$. Che relazione c'è tra lo spettro di \mathcal{T} e quello di \mathcal{T}_N ? E tra lo spettro di \mathcal{X} e quello di $\mathcal{X}_N = \frac{1}{2}(\mathcal{T}_N + \mathcal{T}_N^\dagger)$?

PROB 36.^{#*} Sia $\mathcal{T}_0 = \mathbb{1}$; $\mathcal{T}_1 = \mathcal{T}$ (come definito al problema precedente). Si definisca \mathcal{T}_n in modo che soddisfi la relazione di ricorrenza

$$2\mathcal{T} \mathcal{T}_n = \mathcal{T}_{n+1} + \mathcal{T}_{n-1}$$

Trovare un'espressione esplicita per \mathcal{T}_n per qualunque n . Sia poi \mathcal{N} l'operatore diagonale $\mathcal{N}|n\rangle = n|n\rangle$. Dimostrare che lo spettro dell'operatore autoaggiunto

$$\mathcal{Y} = \mathcal{N}^{-1/2} \frac{\mathcal{T} + \mathcal{T}^\dagger}{2} \mathcal{N}^{-1/2}$$

è dato dai reciproci degli zeri della funzione di Bessel J_0 : $\mathbf{P}\sigma(\mathcal{Y}) = \{y \in \mathbb{R} \mid J_0(1/y) = 0\}$.

PROB 37. Con le notazioni del problema precedente, si determini lo spettro dell'operatore

$$\mathcal{Y}_s = (\mathcal{N} + s\mathbb{1})^{-1/2} \frac{\mathcal{T} + \mathcal{T}^\dagger}{2} (\mathcal{N} + s\mathbb{1})^{-1/2}$$

(s reale non-negativo).

PROB 38. # Nella base ortonormale $|n\rangle$, ($n = 1, 2, 3, \dots$) si definisce l'operatore

$$\mathcal{S}_a |n\rangle = \sum_{j=0}^n a^{j-1} |n-j\rangle$$

con a un qualunque numero complesso di modulo inferiore a uno. Si dimostri che \mathcal{S}_a è limitato, se ne determini lo spettro e si trovi \mathcal{S}_a^\dagger .

PROB 39. # Determinare lo spettro dell'operatore integrale \mathcal{K}_a definito dalla relazione

$$(\mathcal{K}_a f)(x) = \frac{1}{2} \int_{-a}^a e^{-|x-y|} f(y) dy.$$

Generalizzare al caso di un operatore della forma

$$(\mathcal{K} f)(x) = \int_{-a}^a F(|x-y|) f(y) dy.$$

PROB 40. Sia \mathcal{A} l'operatore definito in una base ortonormale dalla matrice

$$\mathcal{A} \leftrightarrow \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \dots \\ a_1 & a_0 & a_1 & a_2 & \dots \\ a_2 & a_1 & a_0 & a_1 & \dots \\ a_3 & a_2 & a_1 & a_0 & \dots \\ \vdots & & & \ddots & \ddots \end{pmatrix}$$

ossia tale che $A_{ij} = a_{|i-j|}$. Determinare se \mathcal{A} è limitato, utilizzando il criterio di Schur.

PROB 41. # Sia p come definito nel Probl.22. Considerare l'operatore $\mathcal{B} = g(x)^{-1} p g(x)$ con $g(x) = x e^{-x^2/2}$. Dimostrare che

$$(\mathcal{B} \mathcal{B}^\dagger f)(x) = -f''(x) + x^2 f(x) - 3f(x).$$

Dunque $\mathcal{B} \mathcal{B}^\dagger e^{-x^2/2} = -2 e^{-x^2/2}$. **Attenzione!** $\mathcal{B} \mathcal{B}^\dagger$ è autoaggiunto e positivo, come può avere un autovalore negativo???

PROB 42. Si dimostri che sotto opportune condizioni su \mathcal{A} e \mathcal{B} l'operatore

$$\mathcal{Y} = \frac{1}{2\pi i} \oint_{\Gamma} (\mathcal{B} - x\mathbb{1})^{-1} X (\mathcal{A} - x\mathbb{1})^{-1} dz$$

soddisfa l'equazione $\mathcal{B} \mathcal{Y} - \mathcal{Y} \mathcal{A} = X$ se il contorno Γ è una curva semplice chiusa contenente al suo interno lo spettro di \mathcal{A} ma nessun punto dello spettro di \mathcal{B} .

PROB 43. Sia $A_{ij} = (1 + i + j)^{-1}$, ($i, j \in \mathbb{Z}^+$). Si dimostri che A definisce un operatore limitato. Si applichi il criterio di Schur, oppure si tenga conto del fatto che A è la matrice dei momenti della distribuzione uniforme nell'intervallo $(0, 1)$

$$A_{ij} = \int_0^1 x^{i+j} dx = \langle x^i, x^j \rangle$$

PROB 44. Siano \mathfrak{p} e \mathfrak{q} gli operatori canonici in $L_2(-\pi, \pi)$, con condizioni periodiche al contorno. Determinare

$$\mathcal{C}(s, t) = \exp\{it\mathfrak{p}\} \exp\{is\mathfrak{q}\} \exp\{-it\mathfrak{p}\} \exp\{-is\mathfrak{q}\}$$

PROB 45. Sia \mathcal{M} la matrice simmetrica

$$\mathcal{M} \leftrightarrow \begin{pmatrix} a_1 & b_1 & 0 & 0 & 0 & 0 & \dots \\ b_1 & a_2 & b_2 & 0 & 0 & 0 & \dots \\ 0 & b_2 & a_1 & b_1 & 0 & 0 & \dots \\ 0 & 0 & b_1 & a_2 & b_2 & 0 & \dots \\ 0 & 0 & 0 & b_2 & a_1 & b_1 & \dots \\ \vdots & \vdots & & & \ddots & \ddots & \ddots \end{pmatrix}$$

($\mathcal{M}_{2n-1, 2n-1} = a_1$, $\mathcal{M}_{2n, 2n} = a_2$, $\mathcal{M}_{2n-1, 2n} = b_1$, $\mathcal{M}_{2n, 2n+1} = b_2$). Determinare lo spettro di \mathcal{M} (assumendo che la base sia ortonormale).

Soluzioni

#5 Dobbiamo chiederci in quale circostanze le infinite relazioni $\langle e_j | \xi \rangle = 0$ implichi $|\xi\rangle = 0$: dato che $\langle e_j | \xi \rangle = a \xi_n + b \xi_{n+1}$ si trova che per $|a| < |b|$ esiste il vettore

$$(1, -b/a, (-b/a)^2, \dots, (-b/a)^n, \dots) \in \ell_2$$

e ortogonale a tutti i vettori della successione, la quale perciò non costituisce una base.

#6 È sufficiente considerare la funzione

$$f(x) = \sqrt{\sum_{n=1}^{\infty} g(n^2(x-n))}$$

con $g(x) = 0$ per $x > 1/2$, $\int g(x) dx = 1$, g continua e positiva con tutte le derivate continue. Si trova

$$\int f(x)^2 dx = \sum_{n=1}^{\infty} \int g(n^2(x-n)) dx = \sum n^{-2} = \zeta(2) = \pi^2/6$$

Dunque f è in L_2 ma non ha limite per $x \rightarrow \infty$. Se si assume che la funzione sia in L_2 , sia assolutamente continua e con derivata in L_2 segue invece:

$$f(x)^2 = f(0)^2 + \int_0^x \frac{d}{dy} f(y)^2 dy = f(0)^2 + 2 \int_0^x f(y) f'(y) dy$$

e il limite di f all'infinito esiste per la convergenza del prodotto interno di due funzioni di L_2 . Inoltre il limite, se esiste, è ovviamente uguale a zero.

#7 La relazione che definisce l'aggiunto

$$\langle \chi, \mathfrak{G} \phi \rangle = \langle \eta, \phi \rangle$$

per $\phi = |1\rangle, |2\rangle, \dots, |n\rangle, \dots$ mostra che

$$\langle \chi | j \rangle = \langle \chi | k \rangle$$

se $\mathfrak{G} |j\rangle = \mathfrak{G} |k\rangle$. Il punto è che esistono infiniti k per ogni j che soddisfano la condizione e quindi $\langle j | \chi \rangle$ assume lo stesso valore infinite volte e deve perciò annullarsi. Questo vale per ogni j e dunque $|\chi\rangle = 0$ e \mathfrak{G}^\dagger ha dominio vuoto. \mathfrak{G} è ovviamente illimitato, altrimenti si contravverrebbe al teorema di Riesz, tuttavia ciò si può controllare direttamente: si ha

$$\psi = |1\rangle + |2\rangle + |4\rangle + \dots + |k_n\rangle; \quad \|\mathfrak{G}\psi\| = n = \sqrt{n} \|\psi\|$$

e dunque $\|\mathfrak{G}\psi\| / \|\psi\|$ può essere grande a piacere.

#9 Il problema si può mappare su uno già risolto: infatti si consideri la trasformazione

$$f(x) = x g(x)$$

Dato che $0 < a < x < b$ la trasformazione è continua e invertibile. Si avrà dunque l'equazione agli autovalori

$$\frac{i}{2} \left(x \frac{d}{dx} x^{-1} + x^{-1} \frac{d}{dx} x \right) g(x) = \lambda g(x)$$

ma si ha anche

$$x \frac{d}{dx} x^{-1} + x^{-1} \frac{d}{dx} x = x(x^{-1} \frac{d}{dx} - x^{-2}) + x^{-1}(x \frac{d}{dx} + 1) = 2 \frac{d}{dx}$$

e l'equazione si riduce a $i \frac{d}{dx} g(x) = \lambda g(x)$ con condizioni al contorno $g(a) = g(b)$, dunque $g(x) = \exp\{i k x\}$ con autovalori $\{k | (b-a)k = 2n\pi, n \in \mathbb{Z}\}$.

#10 Dalla relazione

$$\int_{-\infty}^{\infty} g(x) \frac{d}{dx} f(x) e^{-x^2} dx = - \int_{-\infty}^{\infty} f(x) (e^{x^2} \frac{d}{dx} g(x) e^{-x^2}) e^{-x^2} dx$$

si conclude che

$$\mathcal{B}^\dagger g(x) = -e^{x^2} \frac{d}{dx} e^{-x^2} g(x) = (-\frac{d}{dx} + 2x) g(x)$$

Lo spettro di \mathcal{B} è puntuale e riempie tutto il piano complesso, infatti

$$\mathcal{B}f(x) = f'(x) = b f(x) \Rightarrow f(x) = \mathcal{N} e^{bx} \in \mathfrak{H}$$

mentre lo spettro di $\mathcal{B}^\dagger \mathcal{B}$ si determina, secondo quanto si sa dalla Meccanica Quantistica, in base alla relazione di commutazione

$$\mathcal{B} \mathcal{B}^\dagger - \mathcal{B}^\dagger \mathcal{B} = [\frac{d}{dx}, 2x - \frac{d}{dx}] = 2\mathbb{1}.$$

Ne segue infatti che \mathcal{B} è unitariamente equivalente a un multiplo dell'operatore di annichilazione $(p - iq)/\sqrt{2}$. Lo spettro di $\mathcal{B}^\dagger \mathcal{B}$ è perciò quello dell'oscillatore armonico - si veda qualunque testo di Meccanica Quantistica.

#11 Applicare la formula di Baker–Campbell–Hausdorff, confidando che non ci siano difficoltà nascoste nel carattere non-limitato degli operatori. Si può anche verificare che l'identità è corretta cercando l'equazione differenziale (in t) soddisfatta da entrambi i membri dell'equazione. Una volta assodata la validità della formula, il risolvete è dato dalla Eq.4.10 e l'integrale è assolutamente convergente per ogni valore complesso di z ; dunque $\text{RES}(\mathcal{A}) = \mathbb{C}$.

⚠ La terminologia in letteratura può indurre in errore. Halmos [Hal82] presenta un teorema secondo cui ogni operatore nello spazio di Hilbert ha almeno un punto nello spettro. Il fatto è che nel libro citato si considerano solo operatori definiti su tutto lo spazio di Hilbert, ad es. $\text{diag}\{0, 1, 2, \dots, n, \dots\}$ **non** è la matrice di un operatore. L'argomento portato da Halmos considera il generico elemento di matrice del risolvete $A_{fg}(z) = \langle f, (z\mathbb{1} - \mathcal{A})^{-1} g \rangle$ e fa uso del fatto che $A_{fg}(z) \rightarrow 0$ per $z \rightarrow \infty$ per concludere che lo spettro non può essere vuoto in forza del teorema di Liouville (argomento analogo a quello che si richiama nel caso di spazi a numero finito di dimensioni e che porta al teorema fondamentale sugli zeri dei polinomi in campo complesso). Ma l'argomento non vale per operatori illimitati. ◀

#12 Applicare il criterio di Schur assumendo $p(y) = 1/\sqrt{y}$, $q(y) = 1$. L'inversione dell'operatore risale ad Abel: si applica \mathcal{V} due volte e scambiando l'ordine di integrazione si trova

$$(\mathcal{V}^2 f)(x) = \int_0^x f(y) dy$$

e quindi $\frac{d}{dx} \mathcal{V}^2 = \mathbb{1}$ (vedi [Tri85]).

#13 Si trova una formula di inversione in modo analogo al problema precedente:

$$\mathcal{V}_\alpha \mathcal{V}_{1-\alpha} f(x) = (\sin(\pi\alpha))^{-1} \int_0^x f(y) dy$$

Si applichi poi il criterio di Schur.

#14 La funzione di Green $G_z(x, y) = \langle x | (z\mathbb{1} - p^2) | y \rangle$ si trova risolvendo l'equazione differenziale

$$z G_z(x, y) + \frac{\partial^2}{\partial x^2} G_z(x, y) = \delta(x - y)$$

tenendo conto delle condizioni al contorno; la soluzione $A_{\geq} \cos(\sqrt{z}x + \delta_{\geq})$ con costanti differenti per $x \gtrless y$ si trova imponendo che

$$\int_{y-\varepsilon}^{y+\varepsilon} G_z'' dx = G_z'(y + \varepsilon, y) - G_z'(y - \varepsilon, y) \approx 1$$

(discontinuità nella derivata prima). Si trova, ponendo $z = \omega^2$.

$$G_z(x, y) = \frac{1}{2\omega \sin(\omega\pi)} \cos(\omega(\pi - |x - y|))$$

Una semplice istruzione in *mathematica* risolve il problema

```
DSolve[{z f[x] + f''[x] == DiracDelta[x - y], f[-pi] == f[pi], f'[-pi] == f'[pi]}, f[x], x];
Simplify[%, {z > 0, y > -pi, y < pi, x > y}]
```

#16 Le funzioni u_n sono autofunzioni dell'operatore \mathcal{F} (la trasformata di Fourier - 4.3.8) con autovalore $e^{-in\pi/2}$. Dato che $p = \mathcal{F} q \mathcal{F}^\dagger$ si avrà

$$\langle u_n, (q + \mathcal{F} q \mathcal{F}^\dagger) u_m \rangle = (1 + e^{i(m-n)\pi/2}) \langle u_n, q u_m \rangle$$

da cui discendono la seconda e la terza relazione, mentre la prima è conseguenza della parità $u_n(-x) = (-)^n u_n(x)$. Gli integrali che permettono di costruire la matrice

$$\int_0^\infty u_n u_m x e^{-x^2/2} dx$$

sono esprimibili in termini di funzioni Ipergeometriche, ma sono più comodamente calcolati con *mathematica*. Il calcolo dello spettro si può impostare prendendo un troncamento della matrice a dimensione finita. Tuttavia la matrice richiede un consistente tempo di elaborazione e non è pratico raggiungere dimensioni rilevanti (≥ 64). Un'altra strategia di calcolo è chiaramente superiore: si utilizza la rappresentazione spettrale per $|p\rangle$ e l'algoritmo di Arnoldi per il calcolo dello spettro; 0.2sec per il calcolo con una discretizzazione di 1024 punti. Il codice matlab è il seguente (depurato da qualche dettaglio tecnico, il codice completo si trova in rete)

```

                                absp.m
function [E,Psi,x] = absp(N, L, Neig)
% Spectral code: |p| in momentum space
% Usage:
%   [E,psi,x]=abs(N,L,Neig)
% .....
%% space Lattice
dx = 2*L/N;
x = -L*(1-1/N) : dx : L*(1-1/N);
r = abs(x');

%% lattice in momentum space
kx = fftshift(pi*(-N/2:N/2-1)/L)';
k = abs(kx);
[Psi,E] = eigs(@Op, N, Neig, 'SA');
E = diag(E);

%% embedded routine:
%% compute the action of Operator over |psi>
function Hpsi = Op(psi)
    Hpsi = r .* psi + real(ifft(k.*fft(psi)));
end % end embedded routine
end % end main routine

```

Si può stimare lo spettro attraverso l'approssimazione WKB secondo cui detto E_n l' n -esimo autovalore si ha

$$\frac{1}{2\pi} \iint_{|q|+|p|=E_n} dq dp \approx (n + 1/2)$$

L'integrale è ovviamente dato da $2E_n^2$ e quindi si avrà approssimativamente

$$E_n \approx \sqrt{\pi(n + 1/2)}, n = 0, 1, 2, \dots$$

Il grado di approssimazione della formula WKB si può stimare confrontandola con il valore numerico ottenuto con `matlab`. L'errore $1 - E_{\text{WKB}}/E$ sull'autovalore n -esimo risulta dell'ordine di $10^{-3}/n$ ($n \gg 1$).

#18 Dato che $e^{-tA} \mathcal{B} e^{tA} = \mathcal{B} - tC$ e inoltre e^{-tA} è limitato e quindi costituisce una trasformazione di similitudine, lo spettro di \mathcal{B} coincide con quello di $\mathcal{B} - tC$, ovvero lo spettro di \mathcal{B}/t con quello di $-C + \mathcal{B}/t$. Se prendiamo t molto grande il raggio spettrale di \mathcal{B}/t diventa piccolo a piacere e quindi ciò vale anche per C .

#20 Introduciamo la base ortonormale

$$u_n(z) = \mathcal{N}_n z^n$$

Si trova facilmente

$$\|u_n\|^2 = \iint_{|z|<1} |z|^{2n} dz d\bar{z} = \frac{\pi}{n+1}$$

La serie

$$K(\bar{w}, z) = \sum_{n=0}^{\infty} \overline{u_n(w)} u_n(z) = \frac{1}{\pi} (1 - z \bar{w})^{-2}$$

definisce la funzione $e_w(z) = K(\bar{w}, z) \in H_2(\mathbb{D})$ richiesta, tale cioè che

$$\langle e_z, f \rangle = f(z)$$

#21

- (a) In modo analogo al problema precedente, si vede subito che introdotta la base ortonormale

$$u_n(z) = (n!)^{-1/2} z^n$$

si trova

$$K(\bar{z}, w) = \sum_{n=0}^{\infty} \overline{u_n(z)} u_n(w) = \exp\{\bar{z} w\}$$

e perciò

$$\delta_z(f) = f(z) = \langle e_z, f \rangle$$

dove

$$e_z(w) = K(\bar{z}, w) .$$

Il risultato discende anche in modo molto limpido dal seguente fatto: lo spazio di Bargmann origina da $L_2(-\infty, \infty)$ introducendo la base degli stati coerenti

$$|z\rangle = \sum_{n \geq 0} \frac{\bar{z}^n}{\sqrt{n!}} |n\rangle$$

essendo $|n\rangle$ la base ortonormale costituita dagli autostati dell'oscillatore armonico. La componente di un vettore qualunque $|f\rangle$ lungo $|z\rangle$ è una funzione intera in \mathfrak{B} e per la disuguaglianza di Cauchy si ha

$$|f(z)|^2 = |\langle z|f\rangle|^2 = \left| \sum_{n \geq 0} \frac{z^n}{\sqrt{n!}} \langle n|f\rangle \right|^2 \leq \sum \frac{|z|^{2n}}{n!} \cdot \|f\|^2 = e^{|z|^2} \|f\|^2$$

- (b) L'aggiunto di d/dz si trova dalla definizione: integrando per parti si trova

$$\iint e^{-|z|^2} \overline{f(z)} \frac{dg(z)}{dz} dz \wedge \bar{dz} = \iint \frac{\partial e^{-|z|^2} \overline{f(z)}}{\partial z} g(z) dz \wedge \bar{dz} = \iint e^{-|z|^2} \overline{z f(z)} g(z) dz \wedge \bar{dz}$$

$$\text{e dunque } \left(\frac{d}{dz} \right)^\dagger f(z) = z f(z).$$

- (c) Lo spettro dell'operatore \mathcal{H} si trova facilmente risolvendo un'equazione differenziale del primo ordine:

$$(z + c) \frac{df}{dz} = (\lambda - \bar{c}) f(z)$$

che ha soluzione generale

$$f(z) = \mathcal{N} e^{-\bar{c}z} (z + c)^{\lambda + |c|^2}$$

Affinché f sia una funzione intera è necessario che $\lambda + |c|^2 \in \mathbb{Z}^+$, ossia lo spettro è dato da $\lambda = n - |c|^2$, $n = 0, 1, 2, 3, \dots$

- (d) Si ha

$$a e^{i\omega t} = e^{-i\omega t} a^\dagger a a e^{i\omega t} a^\dagger a$$

e perciò

$$i \frac{d\psi}{dt} = \mathcal{H}_t \psi, \quad \mathcal{H}_t = e^{-i\omega t} a^\dagger a (a^\dagger a + c a + \bar{c} a^\dagger) e^{i\omega t} a^\dagger a$$

Si applica la "rappresentazione di interazione" (a Pag. 67) per ottenere

$$\psi(t) = e^{-i\omega t} a^\dagger a e^{-i\omega t (\mathcal{H}_0 - \omega a^\dagger a)} \psi(0)$$

ed il problema è ricondotto al calcolo di esponenziali di operatori non dipendenti esplicitamente dal tempo. In casi analoghi non esattamente solubili si può ricorrere all'integrazione numerica. Il problema si può mappare su L_2 e l'equazione di evoluzione diventa un'ordinaria equazione di Schroedinger con potenziali dipendenti dal tempo risolvibile con routines standard.

(e) Soluzione analoga al punto precedente.

(f) Si consideri \mathcal{V}_t mappato su L_2 e si tenga conto del fatto che e^{t^a} si può riesprimere in termini di $e^{tq/\sqrt{2}}$. Questo operatore non può essere applicato a qualunque elemento di L_2 a meno che t non sia puramente immaginario, ma allora il problema rinasce dall'altro fattore $e^{itq/\sqrt{2}}$!

#22 Il vettore $|z\rangle$ autovettore di \mathcal{A} con autovalore \bar{z} si trova risolvendo un'equazione differenziale del primo ordine senza difficoltà e si trova

$$\langle x|z\rangle = \mathcal{N} \exp\left\{\sqrt{2}x\bar{z} - \frac{1}{2}(x^2 + \bar{z}^2) - \frac{1}{2}|z|^2\right\}$$

Il prodotto $\langle z|f\rangle$ definisce una funzione in \mathfrak{B} a meno del fattore $e^{-|z|^2/2}$ che viene associato alla misura di integrazione:

$$\langle z|f\rangle = \mathcal{N} \int_{-\infty}^{\infty} \exp\left\{\sqrt{2}x\bar{z} - \frac{1}{2}x^2\right\} f(x) dx \exp\left\{-\frac{1}{2}(z^2 + |z|^2)\right\}$$

La trasformazione cercata è perciò

$$(\mathcal{K}f)(z) = \mathcal{N} \int_{-\infty}^{\infty} \exp\left\{\sqrt{2}x\bar{z} - \frac{1}{2}x^2 - \frac{1}{2}z^2\right\} f(x) dx$$

e si verifica poi che

$$\frac{d}{dz} \mathcal{K} = \mathcal{K} (q + ip)/\sqrt{2}.$$

#23 Si veda l'articolo sul Nuovo Cim.65A, pag.298 (1981), e anche [AO84].

#24 Come per i Probl.19-20-21 si determina una base ortonormale

$$e_n(z) = \mathcal{N}_n z^n$$

e si costruisce il nucleo $K(\bar{z}, w) = \sum \overline{e_n(z)} e_n(w)$ che costituisce la “delta di Dirac” in questi spazi di funzioni analitiche. Si trova $\mathcal{N}_n = \sqrt{\Gamma(n+2\ell)/\Gamma(2\ell)n!}$ e infine

$$K(\bar{z}, w) = (1 - \bar{z}w)^{-2\ell}.$$

Il punto iv) si discute operando una trasformazione di variabile

$$w = \frac{az + b}{\bar{b}z + \bar{a}}$$

e constatando che per $s = \ell$ lo Jacobiano cancella esattamente i fattori davanti a $f(z)$ e quindi \mathcal{U} risulta unitario. In questo modo si costruiscono le rappresentazioni del gruppo di Lorentz in due dimensioni $SO(2, 1) \sim SU(1, 1)$ [Bar47].

#25 Si vedano [VW06, BCO07] per la soluzione analitica. Come esercizio di analisi numerica, si può invece affrontare il problema molto facilmente utilizzando un programma `matlab`. La matrice infinita si deve troncare a dimensione finita N ; per ogni scelta di N i primi $r(N)$ autovalori saranno calcolati con precisione accettabile, $r(N)$ essendo da valutare caso per caso: il principio cui affidarsi in mancanza d'altre informazioni è che gli autovalori sono affetti da un errore di taglia finita, analogo a quanto succede ponendo un sistema quanto-meccanico in una scatola finita: la presenza di un volume finito in genere aumenta il valore dell'energia (si pensi a una “compressione adiabatica” dal volume infinito a L). L'errore sarà meno rilevante per gli autostati il cui autovettore è ampiamente contenuto nel volume finito. Ad es. per un oscillatore armonico di cui si sa che la scala di lunghezze tipiche è fissata da $\sqrt{\hbar/m\omega}$ e la funzione d'onda è in gran parte interna alla zona “classica” $\frac{1}{2}m\omega^2 x^2 \simeq n\hbar\omega$ è facile determinare la dimensione della scatola di normalizzazione. In un problema come quello in esame, non c'è intuizione fisica che aiuti, tuttavia si può procedere allargando via via la dimensione della matrice fino a raggiungere stabilità. Per affrontare un problema agli autovalori di grandi dimensioni si può

utilizzare una routine *matlab* come *eigs* che calcola lo spettro di matrici sparse. Nel nostro caso si definisce la matrice attraverso le routines *spdiags* che producono una matrice sparsa da sottoporre poi a *eigs*.

```

function [E,V,H] = susyqm(lam, B, Negv, flag)
% Veneziano-Wosiek susyQM
% compute the spectrum through sparse methods (eigs) for large
% matrices, dim=B
% Usage [E,V,H] = susyqm(lam, B, Negv)
% ----- omesso il setup, vedi file susyqm.m -----
% building the sparse matrices H
D = (0:B)'; % diagonal m.e. (column vector)
OD = sqrt(lam*D.*(D+1)); % off-diagonal m.e.

H = spdiags([OD,D*(1+lam)/2], [-1:0], B, B); % subdiagonal part
H = H+H'; % symmetrize
H(2,2) = 1 + flag * lam; % fix special value at n = 1
[V,E] = eigs(H(2:end,2:end), Negv, 'SM',opt);

```

Si consiglia di sperimentare il codice per vari valori di λ . Si troverà che per $\lambda \ll 1$ un troncamento 100×100 è già adeguato, mentre per λ crescente e dell'ordine di uno il problema tende a scappare di mano. In effetti per $\lambda = 1$ l'operatore attraversa una transizione a un diverso regime; per $\lambda = 1$ lo spettro è continuo. Ci si accorge facilmente di quanto accade intorno a $\lambda = 1$ in base al fatto che gli autovalori diventano fortemente dipendenti dal troncamento finito. Ad es. per $\lambda = .9999$ lo spettro calcolato per $B = 10^n$, $n = 3, 4, 5, 6$ mostra che solo per dimensione di 10^6 si intravede stabilità rispetto al troncamento (il calcolo prende circa un minuto). La dimensione massima raggiungibile dipende dalle caratteristiche dell'hardware e da quanto è sparsa la matrice. Nel caso presente $Negv = 10, B = 4 \times 10^6$ è già vicino al limite in un sistema con 4GBy di memoria fisica.

dipendenza dal troncamento				
n\B	10^3	10^4	10^5	10^6
1	0.0002	0.0000	0.0000	0.0000
2	0.0049	0.0005	0.0001	0.0001
3	0.0150	0.0014	0.0002	0.0002
4	0.0303	0.0029	0.0004	0.0003
5	0.0508	0.0048	0.0006	0.0004
6	0.0765	0.0073	0.0008	0.0005
7	0.1074	0.0103	0.0011	0.0006
8	0.1435	0.0138	0.0014	0.0007
9	0.1848	0.0178	0.0018	0.0008
10	0.2313	0.0222	0.0023	0.0009

#28 Si usi la trasformata di Fourier.

#29 Si veda la soluzione completa su [CMO80]. L'operatore troncato a dimensione finita può essere agevolmente studiato numericamente. Si veda il Probl.50 nella Parte II.

#30 Si veda [FDO04] per la soluzione.

#31 Iniziamo a dimostrare che, per ragioni puramente analitiche, il sottospazio \mathcal{P}_n dei polinomi di grado n è invariante sotto l'azione di \mathcal{Z} (tutti gli integrali sono estesi a $(-1, 1)$):

$$\begin{aligned} \frac{1}{\pi} \int \sqrt{1-y^2} \frac{y^n}{x-y} dy &= \frac{1}{\pi} \int \sqrt{1-y^2} \frac{y^n - x^n}{x-y} dy + x^n \frac{1}{\pi} \int \frac{\sqrt{1-y^2}}{x-y} dy \\ &= - \sum_{k=0}^{n-1} x^k \frac{1}{\pi} \int \sqrt{1-y^2} y^{n-1-k} dy + x^{n+1} \\ &= x^{n+1} - x^{n-1}/2 + O(x^{n-3}) \end{aligned}$$

e perciò

$$\mathcal{Z} x^n = (n+1)x^n - \frac{1}{2}(n-1)x^{n-2} + O(x^{n-4})$$

Nella base dei monomi $\{x^k\}$ l'operatore \mathcal{Z} è quindi rappresentato da una matrice triangolare superiore

$$\mathcal{Z} \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & \dots & \dots & \dots & \dots \\ 0 & 2 & 0 & 0 & \dots & \dots & \dots \\ 0 & 0 & 3 & 0 & -1/2 & \dots & \dots \\ 0 & 0 & 0 & 4 & 0 & -1 & \dots \\ \dots & \dots & \ddots & \ddots & \ddots & \ddots & \dots \\ 0 & 0 & 0 & \dots & 0 & n & 0 & 1-\frac{1}{2}n & \dots \\ \dots & \dots & \dots & \ddots & \ddots & \ddots & \ddots & \ddots & \dots \end{pmatrix}$$

e di conseguenza \mathcal{Z} ha per autovalori gli interi positivi. Nel sottospazio invariante \mathcal{P}_n ci sono $n+1$ autovettori; tra questi uno, quello appartenente all'autovalore $n+1$, è ortogonale a \mathcal{P}_{n-1} rispetto alla metrica

$$(4.15) \quad \langle g|f \rangle = \int_{-1}^1 \overline{g(x)} f(x) \sqrt{1-x^2} dx$$

L'operatore \mathcal{Z} è infatti Hermitiano rispetto a questa metrica. Per vederlo, conviene trasformare l'espressione integrale come segue

$$\begin{aligned} \frac{1}{\pi} \frac{d}{dx} \int \sqrt{1-y^2} \frac{f(y)}{x-y} dy &= \frac{1}{\pi} \frac{d}{dx} \int \sqrt{1-y^2} \frac{f(y) - f(x)}{x-y} dy \\ &\quad + \frac{1}{\pi} \frac{d}{dx} \left(f(x) \int \sqrt{1-y^2} \frac{1}{x-y} dy \right) \\ &= \frac{d}{dx} (x f(x)) + \frac{1}{\pi} \int \sqrt{1-y^2} \frac{d}{dx} \frac{f(x) - f(y)}{y-x} dy \\ (4.16) \quad &= f(x) + \frac{1}{\pi} \int \sqrt{1-y^2} \frac{f(x) - f(y)}{(x-y)^2} dy \end{aligned}$$

e in questa seconda formulazione il carattere Hermitiano di \mathcal{Z} è evidente. E a questo punto si è risolto anche il Probl. 29, che è assolutamente equivalente! Le autofunzioni sono polinomi ortogonali nella metrica dell'Eq. 4.15, e dunque si tratta di polinomi di Tchebychev di secondo tipo $U_n(x)$. Essenzialmente l'unico integrale da valutare è

$$\frac{1}{\pi} \int \frac{\sqrt{1-y^2}}{x-y} dy = x$$

che discende immediatamente dall'Eq. (4.16), ed è facilmente verificato con *Mathematica*, utilizzando l'opzione `Integrate[... , PrincipalValue→True]`. Per maggiori dettagli si vedano [Tri85, CMO80, CP78].

#34 Sviluppando in serie di Fourier $\psi(\vartheta) = \sum \psi_n e^{in\vartheta}$, si trova

$$\mathcal{H}[\psi] = \sum_{-\infty}^{\infty} |n| |\psi_n|^2;$$

inoltre la continuità di ψ implica che

$$\nu[\psi] = \frac{\psi(2\pi) - \psi(0)}{2\pi} = \int_0^{2\pi} \left(e^{-i\psi(\vartheta)} (-i) \frac{d}{d\vartheta} e^{i\psi(\vartheta)} \right) \frac{d\vartheta}{2\pi} = \sum_{-\infty}^{\infty} n |\psi_n|^2;$$

è un intero che rimane costante rispetto a piccole variazioni di ψ (rappresenta il numero di avvolgimenti del punto $e^{i\psi}$ intorno allo zero nel piano complesso). Per $\nu > 0$ fissato si ha allora

$$\mathcal{H}[\psi] - \nu[\psi] = \sum_{-\infty}^{\infty} (|n| - n) |\psi_n|^2 = -2 \sum_{-\infty}^{-1} n |\psi_n|^2 \geq 0$$

Il minimo si ha per $\psi_n = 0, n = -1, -2, \dots, -\infty$, il che implica che la funzione $e^{i\psi(\vartheta)} = \sum_0^{\infty} \psi_n z^n$ si estende a una funzione analitica regolare all'interno del disco $|z| = |e^{i\vartheta}| < 1$. Le condizioni

$$|e^{i\psi(\vartheta)}| = 1 \quad \text{per } |z| = 1$$

$$\oint \psi'(z) dz = 2\pi\nu$$

restringono la ψ alla forma

$$e^{i\psi} = e^{i\alpha} \prod_{i=1}^{\nu} \frac{z - z_i}{1 - \bar{z}_i z}$$

dove z_i sono punti arbitrari all'interno del cerchio di raggio uno. Per $\nu < 0$ si procede in modo analogo (vedi [Ono85]).

#35 L'equazione agli autovalori per il vettore $\langle n | \lambda \rangle = c_n$

$$c_{n+1} + c_{n-1} = 2\lambda c_n$$

si può risolvere ponendo $c_n = e^{in\alpha}$ che dà $\lambda = \cos \alpha$. La condizione per $n = 1$ impone di prendere la combinazione di soluzioni che si annulla per $n = 0$, cioè $c_n = \sin(n\alpha)$. La soluzione non è a quadrato sommabile dunque $\mathbf{P}\sigma = \emptyset$, mentre $\mathbf{C}\sigma = [-1, 1]$. Il troncamento a dimensione finita dà lo spettro già studiato nella prima parte, che al limite $N \rightarrow \infty$ riempie l'intervallo $[-1, 1]$; dunque in questo caso c'è un limite regolare verso la dimensione infinita. Al contrario per l'operatore isometrico \mathcal{T} la situazione è nettamente diversa. Lo spettro di \mathcal{T} è dato dall'interno del disco di raggio unitario, mentre il suo troncamento $N \times N$ costituisce un blocco nilpotente con il solo zero nello spettro.

#36 L'equazione agli autovalori $\mathcal{Y}|y\rangle = y|y\rangle$, sviluppata sulla base $|n\rangle$ è data da

$$\sum_{n=1}^{\infty} c_n \left((n+1/2)^{-1/2} |n+1\rangle + (n-1/2)^{-1/2} |n-1\rangle \right) = 2y \sum_{n=1}^{\infty} n^{1/2} c_n |n\rangle$$

avendo posto $\langle n | y \rangle = n^{1/2} c_n$. Si trova allora

$$c_{n-1} + c_{n+1} = 2ny c_n$$

che si riconosce come la relazione di ricorrenza delle funzioni di Bessel; dunque $c_n = A J_n(1/y) + B N_n(1/y)$. C'è da imporre una condizione al contorno e precisamente $c_0 = 0$ e inoltre la serie $\sum n |c_n|^2$ deve risultare convergente. Ora si sa che J_n tende a zero per $n \rightarrow \infty$ (è la trasformata di Fourier di una funzione periodica), mentre la seconda soluzione deve divergere in quanto il Wronskiano deve rimanere costante. Dunque $B = 0$ e $c_0 \propto J_n(1/y) = 0$. Il calcolo si implementa facilmente in `matlab` (vedi `bessel.m`).

#38 Si applichi il criterio di Schur. Lo spettro si trova sfruttando la relazione che sussiste tra \mathcal{S}_a e l'operatore di shift:

$$\mathcal{S}_a = \sum_{j \geq 0} a^j \mathcal{T}^j = (\mathbb{1} - a\mathcal{T})^{-1}$$

Noto lo spettro di \mathcal{T} se ne deduce lo spettro di \mathcal{S}_a . Ora \mathcal{T} è limitato con $\|\mathcal{T}\| = 1$. L'equazione $\mathcal{T}|\lambda\rangle = \lambda|\lambda\rangle$ ha soluzione $|\lambda\rangle = \sum_{n=0}^{\infty} \lambda^n |n\rangle$ per ogni $\lambda \in \mathbb{C}$ con $|\lambda| < 1$. Lo spettro di \mathcal{S}_a è perciò dato da $\{(1 - az)^{-1} \mid |z| < 1\}$. Si noti che per ogni n e ogni a si verifica che

$$\mathcal{S}_a(\mathbb{1} - a\mathcal{T})|n\rangle = |n\rangle$$

ma questo non è sufficiente per concludere che $(\mathbb{1} - a\mathcal{T})$ sia sempre invertibile; ciò vale solo per $|a| < 1$.

#39 Prendendo la derivata rispetto a x due volte si scopre che \mathcal{K} è interpretabile come inverso di un operatore differenziale etc.

#41 Non c'è nessun errore, ma solo una manipolazione da prestigiatore che nasconde il trucco. L'operatore \mathcal{B}^\dagger è definito in un dominio che **non comprende** la Gaussiana. Però l'operatore $\mathcal{B}\mathcal{B}^\dagger$ ammette un'estensione autoaggiunta a un dominio che non richiede l'annullamento in $x = 0$ e però non è più fattorizzabile nel prodotto di un operatore per il suo aggiunto. L'esempio risale a lezioni di J.R. Klauder alla Scuola di Schladming del 1969.

#45 Cenni: si può studiare il problema analogamente a quello del Cap.1 tenendo conto che in questo caso una relazione di ricorrenza coinvolge due successioni di determinanti (indice pari o dispari). Se si vuole esplorare il problema per via numerica si veda il codice `bloch.m`.

```

function [E,M]=bloch(r, neig, N)
% Probl.III-45
% Usage [E,M]=bloch(r, neig, N)
% r = 1 per il probl. nel testo,
% neig = numero di autovalori da calcolare
% N = dim della matrice M
----- omessi dettagli -----
copies=floor(N/(2*r));
a=2*ones(2*r,1) + .125*[ones(r,1);-ones(r,1)];
b=-ones(2*r,1);
A=[];B=[];
for j=1:copies,
    A=[A;a];
    B=[B;b];
end
M=spdiags(B,1,N,N);
M = M+M';
M=M+spdiags(A,0,N,N);
opt.disp=0; % quiet display
E=eigs(M, neig, 'SA', opt);

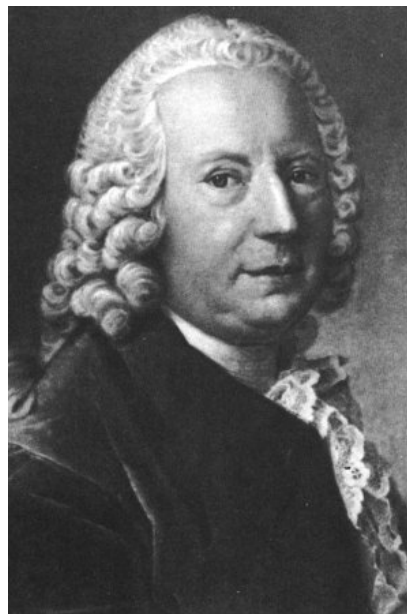
```

Galleria di ritratti

Volti dal mondo della *Matematica*,
della *Fisica* e della *Computer Science*



G. Andrews



D. Bernoulli



Valentine Bargmann



E. Cartan



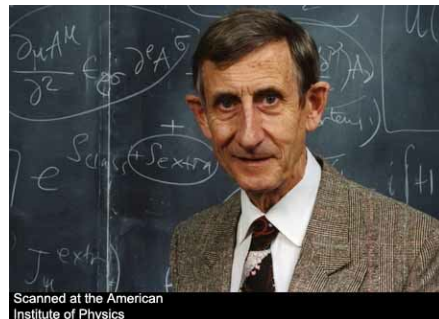
A. Cauchy



L. Dirichlet



P.A.M. Dirac



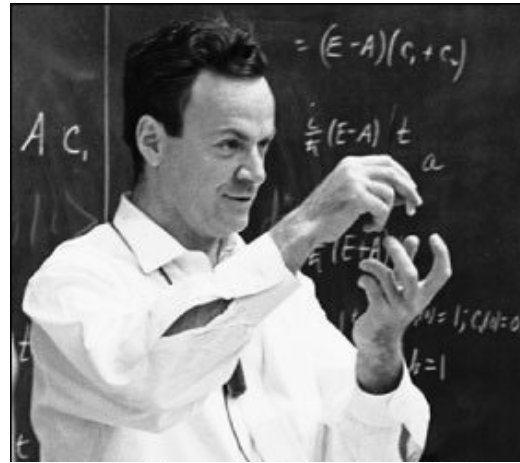
F. Dyson



L. Euler



J. Fourier



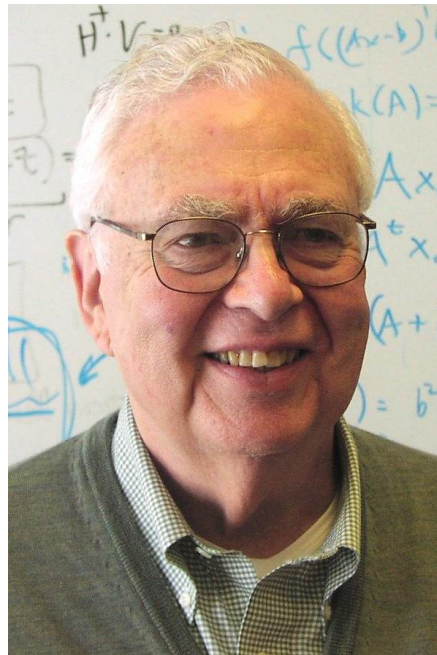
R.P. Feynman



W. Pauli, W. Heisenberg e E. Fermi, Como 1927



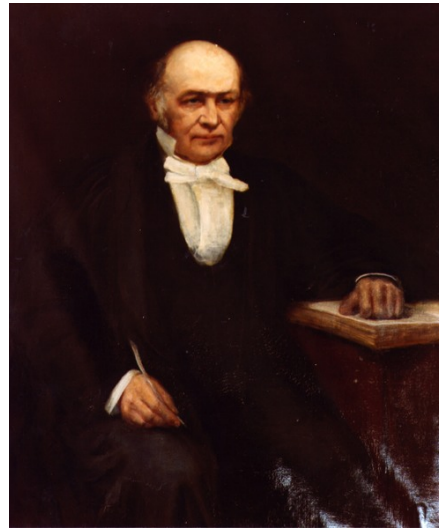
C.F. Gauss



G.H. Golub



G. Hardy



W.R. Hamilton



C. Hermite



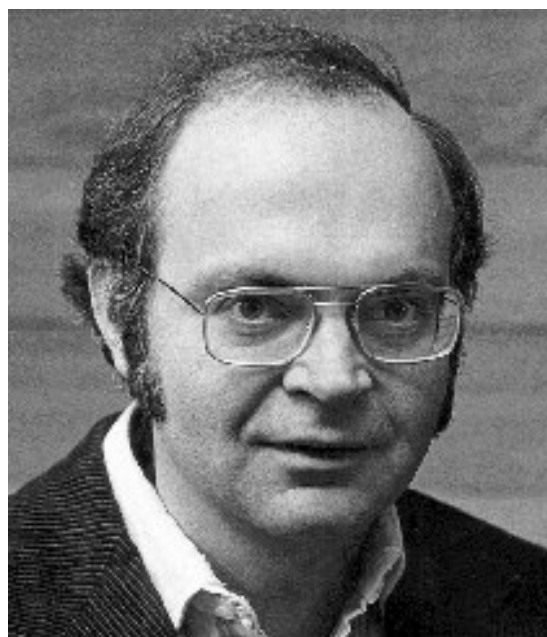
D. Hilbert



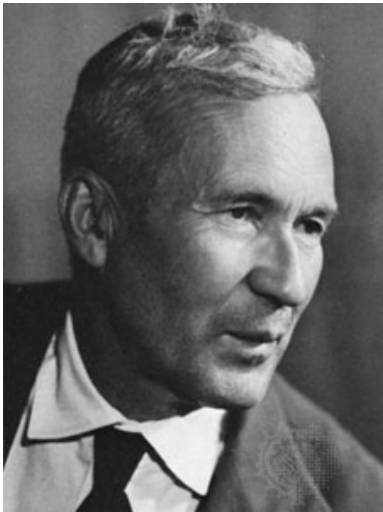
M. Kac



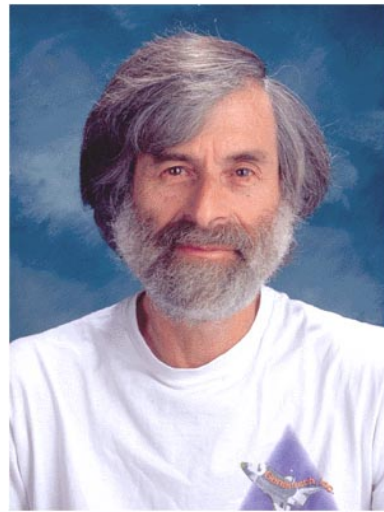
J.R. Klauder



D. Knuth



A.N. Kolmogorov



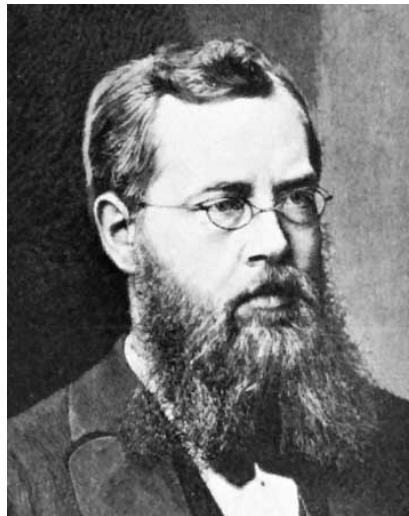
L. Lamport



J.L. Lagrange



A.M. Legendre



S. Lie



A. Markov



M. Mehta



C. Moler



B. Pascal



H. Poincaré



G. Polya



E. Schrödinger



R. Stallman



B. Simon



E.O. Tuck



J. Von Neumann



V. Volterra

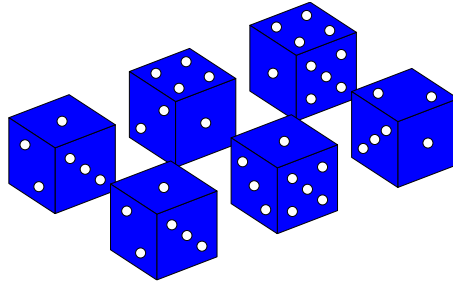


E.P. Wigner

IV - Applicazioni alla Teoria delle Probabilità

CAPITOLO 5

Caso e probabilità in Fisica



Introduzione

In quali circostanze si parla di “caso” e si basano perciò le nostre previsioni su valutazioni di probabilità? Rimanendo nel contesto della Fisica Classica, il caso fa la sua apparizione in fenomeni dove le osservazioni sono dominate da un grande numero di fattori imponderabili, troppo numerosi per potere essere tenuti in conto in modo sistematico; e però proprio il grande numero di fattori permette di prevedere almeno entro certi limiti il risultato delle osservazioni. I casi più comuni e familiari sono dati da fatti legati al gioco: l’uscita di un numero al lotto, il lancio dei dadi, il gioco della roulette etc. In ciascuno di questi casi, il fenomeno che porta alla vincita o alla sconfitta cadrebbe in linea di principio sotto le leggi della meccanica e tuttavia il risultato dipende da un così grande numero di fattori (la velocità iniziale di rotazione della roulette, la velocità e la direzione della pallina, lo stato della superficie della roulette che modifica l’aderenza, l’attrito con l’aria e la presenza di microcorrenti d’aria, le variazioni di temperatura, l’attrito del perno della roulette che determina il tempo di rotazione, cos’altro? Bene, nessuno di questi fattori è alla portata delle misure del giocatore istante per istante (e neppure del croupier) e perciò il massimo che si possa concludere è che per ragioni di pura e semplice simmetria ogni numero della roulette è *a priori ugualmente probabile*. Allo stesso modo l’uscita del lancio di due dadi o i numeri al lotto sono esempi di eventi imprevedibili con sicurezza in quanto determinati da numerosi fattori e da condizioni iniziali non controllabili. Non ci addentreremo nella discussione su quale sia la definizione corretta di probabilità. Spesso si è caduti nella trappola di definire la probabilità in modo logicamente circolare. Perciò adotteremo un punto di vista molto pragmatico. La probabilità P di un evento \mathcal{E} è un numero compreso tra zero e uno e sulla base della conoscenza di questo numero siamo pronti a scommettere che in N esperimenti preparati nelle stesse condizioni l’evento si produrrà approssimativamente NP volte, ossia P è una valutazione a priori della *frequenza* di un evento casuale. L’impostazione di un problema di probabilità è dunque nei casi più semplici fondata su alcuni passaggi fondamentali: i) l’individuazione degli elementi che contribuiscono alla definizione del fenomeno (eventi elementari); l’individuazione tra questi dell’insieme completo di eventi elementari che si possono considerare equiprobabili; iii) l’analisi combinatoria che permette di valutare la probabilità di un evento complesso una volta ridotto ai suoi elementi. Allora, piuttosto che analizzare in modo rigoroso le basi dell’approccio probabilistico (De Finetti, Laplace, Poincaré, Feller) procediamo adottando lo schema matematico della probabilità come ipotesi di lavoro. Le conferme che provengono

dalla sua applicazione a problemi concreti permetterà allora di convincerci della validità dell'approccio. Prima di iniziare un paio di osservazioni che sono per me molto importanti. Prima di tutto il manifestarsi del caso in fenomeni fisici non è legato esclusivamente a quanto detto finora (la presenza di numerosi fattori imponderabili); sappiamo ormai per certo che i fenomeni quantistici sono caratterizzati da comportamenti casuali non riconducibili a semplici modelli di *agenti sub-microscopici e non osservati*. Il caso sembra entrare nelle leggi fisiche in modo intrinseco, anche se la cosa non piaceva al grande Alberto. Gli esperimenti condotti nell'arco degli ultimi trent'anni hanno mostrato che certe misure di correlazione sono incompatibili con una impostazione probabilistica in senso tradizionale (*disuguaglianze di Bell*). Una trattazione particolarmente semplice si trova nell'appendice di W. Faris al libro di D. Wick [Wic55]. Non ci occuperemo di probabilità in ambito quantistico.

Una seconda osservazione è che la ragione per cui un numero sempre crescente di fisici si sono occupati di tecniche probabilistiche sta nel fatto che queste tecniche si rivelano molto potenti e permettono di sviluppare algoritmi di calcolo per la soluzione di problemi che di per sé non hanno nulla di aleatorio*. Ciò è dovuto a particolari coincidenze per cui la soluzione di certe equazioni differenziali (ad es. l'equazione di Laplace) è equivalente alla soluzione di problemi di "camminate aleatorie" (random walk, RW per brevità). In molte situazioni è più agevole simulare un RW piuttosto che risolvere una complicata equazione alle derivate parziali.

Bibliografia ragionata. Normalmente la bibliografia si dà alla fine. Tuttavia è giusto citare subito le fonti per queste lezioni. Mi sono basato sull'agile testo di Rozanov [Roz69], ricco di esempi. Come esercizi iniziali consiglio il libretto di Mosteller [Mos65], con esempi tratti dalla vita di tutti i giorni (più o meno). Un ottimo trattato di riferimento è [PR69]. A un livello più elementare [Gne79]. Inoltre sulla collezione [Wax54] si trovano lavori originali che hanno fatto la storia della applicazione di concetti di probabilità alla teoria fisica (Chandrasekhar, Ornstein-Uhlenbeck, Kac). Il testo di Kac [Kac59] è molto avanzato dal punto di vista matematico e presenta la connessione con la probabilità quantistica. Le problematiche connesse al ruolo della probabilità in fisica quantistica sono presentate in molti testi di meccanica quantistica. Il libro di Bell [Bel87] contiene una raccolta di scritti del fisico irlandese, alla base di molti sviluppi recenti. Ma una presentazione per non specialisti si può trovare utilmente nel libro di Penrose [Pen89], edizione italiana [Pen97]. Il trattato di Feller [Fel68] costituisce il riferimento standard per la teoria delle probabilità e permette di approfondire tutti gli argomenti coperti in queste lezioni. Per le connessioni tra probabilità, random walk e teoria quantistica dei campi si può consultare [ID89], primo volume. Un altro testo ricco di idee, esempi, problemi e che presenta un approccio originale alla teoria della probabilità è il libro di Backlawski, Cerasoli e Rota [BCR84].

Convenzioni adottate nel testo. Utilizziamo i simboli standard di teoria degli insiemi \cup (unione), \cap (intersezione), \emptyset (insieme vuoto). Il *complementare* di un sottoinsieme è indicato con $\complement A = \{\omega \in \Omega | \omega \notin A\}$.

5.1. La definizione matematica di probabilità

La Teoria della Probabilità si è sviluppata come teoria matematica a partire da Pascal e Laplace; successivamente Poincaré e altri matematici del XX secolo hanno dato un impulso notevole e oggi la teoria è una branca della matematica con agganci all'analisi matematica, all'analisi funzionale, alla geometria, alla teoria della misura *etc.* Gli elementi sono molto semplici. Si introduce lo *spazio degli eventi* costituito da un insieme Ω i cui elementi sono detti *eventi elementari*. Ad Ω si applicano le definizioni caratteristiche della teoria degli insiemi, che nella sua forma elementare daremo per scontata. Ad ogni evento elementare $\omega \in \Omega$ si attribuisce

*In una prima versione era scappato un "causale" anziché "casuale", il che induce a utilizzare preferibilmente il termine proveniente dal Latino.

un numero non-negativo $\mathcal{P}(\omega)$ che verrà definito la *probabilità* di ω . In sintesi la definizione assiomatica di probabilità è data da una tripletta di oggetti $\mathcal{P} = \{\Omega, \Sigma, \mathcal{P}\}$, detta *spazio di probabilità*, così specificati:

- i) Ω è un insieme di oggetti detto *lo spazio degli eventi elementari*;
- ii) Σ è una famiglia di sottoinsiemi di Ω , detti *eventi*;
- iii) \mathcal{P} è una *misura* su Σ , ossia una funzione $\mathcal{P} : \Sigma \mapsto [0, 1]$, detta *misura di probabilità*.

L'insieme Ω può essere costituito da un insieme finito (le facce di un dado), infinito numerabile (i punti di un reticolo cristallino), da una varietà geometrica (ad es. un segmento di \mathcal{R} , una regione piana in \mathcal{R}^2 , una qualunque superficie \mathcal{S} , un dominio in \mathcal{R}^n , etc.). I sottoinsiemi che costituiscono Σ devono essere *misurabili*, cioè per ogni $\sigma \in \Sigma$ è definito il valore $0 \leq \mathcal{P}(\sigma) \leq 1$. Vengono assunti alcuni assiomi che caratterizzano uno spazio di probabilità.

- a) Σ è chiusa sotto operazioni di *unione* e *intersezione*, anche in successioni infinite;
- b) $\mathcal{P}(\sigma_1 \cup \sigma_2) = \mathcal{P}(\sigma_1) + \mathcal{P}(\sigma_2)$ se $\sigma_1 \cap \sigma_2 = \emptyset$ (l'insieme vuoto), e più in generale $\mathcal{P}(\bigcup_i \sigma_i) = \sum_i \mathcal{P}(\sigma_i)$ se gli eventi σ_i sono a due a due disgiunti; ($\sigma_i \cap \sigma_j = \emptyset, \forall (i, j)$)
- c) $\mathcal{P}(\emptyset) = 0, \mathcal{P}(\Omega) = 1$.

ESEMPIO 5-1. Si considerino tutti i possibili intervalli contenuti nel segmento $(0, 1)$ e tutte le loro possibili unioni (pluri-intervalli). Ad un qualunque intervallo è associata una lunghezza e questa si può assumere come $\mathcal{P}()$.

PROBLEMA 5-1. Dopo avere mescolato accuratamente un mazzo di 32 carte (composto da $7 : 10, J, Q, K, A$ in 4 semi diversi $\heartsuit \diamondsuit \clubsuit \spadesuit$), si distribuiscono le carte a quattro giocatori di poker. Qual'è la probabilità che il primo giocatore abbia un poker servito?

SOLUZIONE. Ci sono $\binom{32}{5} = 201376$ possibilità per il primo giocatore, (un evento elementare corrisponde a una mano qualunque di 5 carte) e non è consigliabile elencarle qui tutte per contare quante contengono quattro carte uguali; in effetti la cosa si può fare utilizzando un computer: si costruiscono esplicitamente tutte le possibili mani e si conta quante di queste contengono un poker (vedi il programma `matlab`). Tuttavia una enumerazione completa è impensabile già per problemi di poco più complicati. Fortunatamente talvolta si può utilizzare un argomento combinatorio, così com'è per le combinazioni del gioco del poker: in un poker servito la carta diversa dalle altre può essere scelta in 32 modi differenti, e per ogni scelta sono possibili 7 modi per fare poker. Dunque la probabilità, come rapporto tra i casi favorevoli e i casi possibili, è data da $32 \times 7 / \binom{32}{5} = 1/899$, cioè poco più di un caso su mille. E infatti è ritenuto fortunato un giocatore cui si presenti questa circostanza durante la serata. In generale per n carte diverse si ha $\text{prob}(\text{poker}) = 5 / \binom{4n-1}{3}$ (lasciamo la dimostrazione per esercizio). Vogliamo invece ora affrontare il problema attraverso una simulazione, anche se conosciamo già la risposta, giusto per introdurre qualche idea di base.[†] Avendo a disposizione un calcolatore e un linguaggio come `matlab` possiamo simulare un numero arbitrario di smazzate e contare in quanti casi si presenta un poker servito.

[†]Ogni problema di probabilità legato al poker è ampiamente trattato in letteratura. Ad es. si può chiedere a Wikipedia: http://en.wikipedia.org/wiki/Poker_probability.

```

                                poker.m
function p=poker(nC, N)
% Random draws from a deck of nC different cards in 4 suits
% p returns the number of pokers served to the first player
% in N shuffles
% Usage:
% p=poker(nC, N)
..... omitted details .....
C=1:nC;
M=sort([C,C,C,C]);           % 4 suits
p=0;
for j=1:N
    G=M(randperm(4*nC));      % shuffle
    G=sort(G(1:5));
    if (G(1)==G(4) || G(2)==G(5)) % poker!
        p=p+1;
    end;
end

```

Note: M contiene quattro copie identiche (il colore è qui irrilevante) di nC carte; `randperm` al rigo 12 mescola il mazzo, le prime cinque carte, per semplicità vengono date al primo giocatore (contrariamente all'uso, ma agli effetti statistici è irrilevante) e si controlla, una volta messe in ordine crescente, se le prime quattro o le ultime quattro sono uguali, nel qual caso scatta il contatore $p \leftarrow p + 1$. Il risultato permette di valutare oltre il valore atteso (valore medio), anche la fluttuazione del fenomeno, cioè di quanto il numero osservato in ogni simulazione si scosti dal valore medio. La Fig. 5-1 presenta attraverso un istogramma il risultato di 200 partite ognuna con 1000 smazzate. L'istogramma è confrontato con una distribuzione binomiale[‡] con valore medio $1000/899$. Il metodo è denominato “MonteCarlo” e sarà illustrato un po' più in generale nel cap. 5.3.4. Il valor medio calcolato sugli eventi simulati risulta 1.17 non molto lontano da quello esatto (≈ 1.11). Tuttavia si noterà che abbiamo utilizzato un numero di eventi simulati dell'ordine di grandezza dell'intera popolazione. I casi in cui è vantaggioso effettuare una simulazione sono quelli in cui lo spazio degli eventi è molto grande e il campione è scelto in modo mirato ma di dimensioni molto più piccole. ◀

ESEMPIO 5-2. L'insieme Ω sia dato dai giorni dell'anno; la famiglia Σ è data da tutti i possibili sottoinsiemi di giorni dell'anno, ad es. i giorni dispari, oppure le domeniche, etc. Una misura di probabilità è una funzione che associa un numero reale positivo o nullo a ogni giorno dell'anno, con la condizione che la somma dia uno. Ogni particolare fenomeno casuale che sia associato ai giorni dell'anno è associato a una particolare misura. Se ci chiediamo “qual'è la probabilità che in un certo periodo dell'anno la temperatura media superi i 30°C ”, la risposta dipenderà da tanti fattori non controllabili che si dovrà per forza di cose basare la stima sulla statistica degli anni precedenti. Più che di probabilità, in questo caso, si parlerebbe di *inferenza statistica*. Questo genere di problemi non saranno trattati in queste lezioni, per quanto siano della massima importanza in svariati campi. ◀

🔗 Si vedano prima di proseguire i primi esempi nella sezione Problemi a Pag. 213.

L'applicazione dello schema astratto di “spazio di probabilità” a problemi concreti (in Fisica o in qualunque altro contesto) si basa sull'assunto che gli eventi elementari di Ω siano associati ad eventi che si possono osservare e registrare e su cui è possibile effettuare rilevazioni statistiche.

[‡]Diamo per scontato che il lettore conosca le distribuzioni statistiche elementari, Gauss, Cauchy, Bernoulli-binomiale, Poisson. Un richiamo si trova in Appendice.

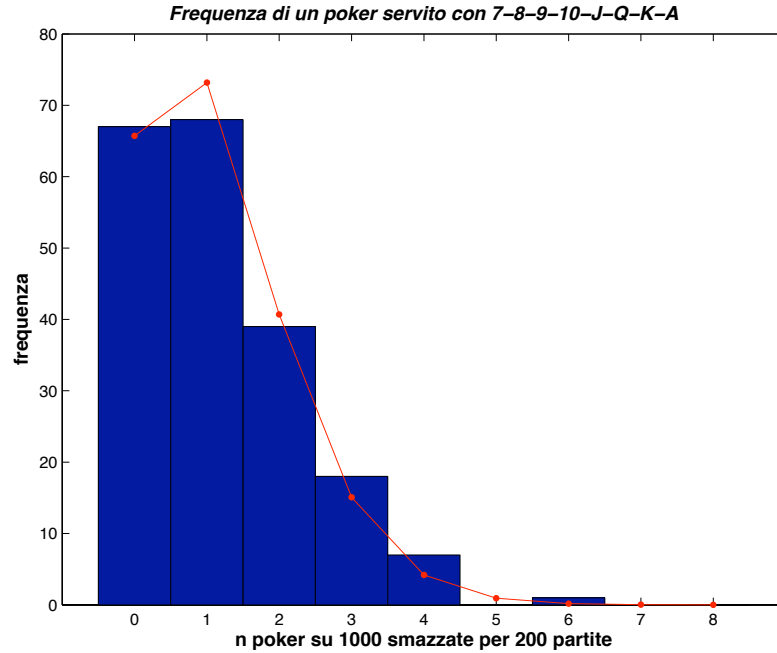


FIGURA 5-1. Una partita a poker simulata.

Se in una serie di osservazioni l'evento ω è osservato $n(\omega)$ volte su un totale di N , il nostro modello probabilistico dovrà essere in grado di prevedere che

$$\frac{n(\omega)}{N} \approx \mathcal{P}(\omega)$$

La questione dell'applicabilità di questi concetti a situazioni reali passa attraverso la precisazione del significato del simbolo \approx . Ci si aspetta che la *frequenza* $\frac{n(\omega)}{n(\Omega)}$ e la probabilità $\mathcal{P}(\omega)$ tendano a coincidere quando il numero di eventi sia sufficientemente grande. Quest'idea si realizza nei vari teoremi "dei grandi numeri". Attenzione però che ci sono pochi concetti quali la *legge dei grandi numeri* ad essere maggiormente travisati dai non esperti, come vicende legate al Lotto dimostrano ampiamente.

5.1.1. La formula di inclusione/esclusione. Siano A_1, A_2, \dots, A_N sottoinsiemi dello spazio degli eventi Ω . Ci si chiede qual'è la probabilità che si realizzi un evento contenuto in uno qualunque dei sottoinsiemi? La risposta $\mathcal{P}(\cup_i A_i) \stackrel{?}{=} \sum_{i=1}^N \mathcal{P}(A_i)$ è ovviamente affetta da errore se gli eventi non sono a due a due disgiunti. Viene spontaneo correggere la formula compensando per le intersezioni

$$(?) \quad \mathcal{P}(\cup_i A_i) \stackrel{?}{=} \sum_i \mathcal{P}(A_i) - \sum_{i>j} \mathcal{P}(A_i \cap A_j)$$

ma anche questa formula non funziona se esistono *triple* intersezioni $A_i \cap A_j \cap A_k$. La risposta corretta è data dalla formula di *inclusione/esclusione* ([Roz69], Teor.2.2, [Poi12], Pag. 59; si veda poi [Sta99] per un punto di vista più ampio).

$$\mathcal{P}(\cup_i A_i) = \sum_i \mathcal{P}(A_i) - \sum_{i>j} \mathcal{P}(A_i \cap A_j) + \sum_{i>j>k} \mathcal{P}(A_i \cap A_j \cap A_k) + \dots - (-1)^n \sum_{i_1>i_2>\dots>i_n} \mathcal{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n})$$

PROVA. Consideriamo un evento elementare che non appartenga ad alcun sottoinsieme A_i : nella formula questo pesa giustamente per zero. Se un evento elementare appartiene a k tra gli A_i , $0 < k \leq n$, secondo la formula esso contribuisce con un peso

$$\binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \dots = - \sum_{j=1}^k (-1)^j \binom{k}{j} = 1.$$

Il risultato risale a Poincaré, [Poi12]. Il “problème de la rencontre” è trattato nel problema 23. ■

60

CHAPITRE III.

D'où, une généralisation du théorème des probabilités totales : la probabilité pour que l'un, au moins, des événements se produise est

$$\Sigma p_i - \Sigma p_{ik} + \Sigma p_{ijk} - \dots$$

23. Problème de la rencontre. — Dans une urne, il y a μ boules numérotées de 1 à μ ; je les tire les unes après les autres, jusqu'à ce que l'urne soit vide. Il y a rencontre si, au i^{e} tirage, je tire la boule numérotée i .

FIGURA 5-2. Estratto da Poincaré, Calcule des Probabilités

5.1.2. Probabilità condizionata. Quanto è influenzato l'evento $A \subset \Omega$ dal verificarsi di un altro evento B ? Ad es. se $B \subset A$ ovviamente A si verifica sempre in seguito a B , mentre se $A \cap B = \emptyset$ i due eventi non possono verificarsi entrambi, e si diranno incompatibili. Si definisce allora la probabilità condizionata di A dato B secondo la formula

$$\mathcal{P}(A|B) \equiv \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}.$$

Supponiamo che la collezione di eventi $\mathfrak{B} = \{B_j | j = 1, \dots, n\}$ sia caratterizzata dal fatto che per ogni coppia $B_j \cap B_k = \emptyset$ e inoltre $\bigcup_{k=1}^n B_k = \Omega$. \mathfrak{B} si dice un insieme completo di eventi e vale la relazione generale

$$(5.1) \quad \mathcal{P}(A) = \sum_{k=1}^n \mathcal{P}(A|B_k) \mathcal{P}(B_k).$$

Si noterà la somiglianza formale con la formula che esprime lo sviluppo di uno stato quantistico su una base

$$|\psi\rangle = \sum_n |n\rangle \langle n|\psi\rangle$$

ma la differenza sta nel fatto che in teoria delle probabilità si compongono i valori di \mathcal{P} , non le ampiezze.

5.1.3. La formula di Bayes. Se B_1, B_2, \dots, B_n è un insieme completo di eventi incompatibili tra loro allora vale la formula

$$\mathcal{P}(B_k|A) = \frac{\mathcal{P}(B_k) \mathcal{P}(A|B_k)}{\mathcal{P}(A)} = \frac{\mathcal{P}(B_k) \mathcal{P}(A|B_k)}{\sum_{k=1}^n \mathcal{P}(B_k) \mathcal{P}(A|B_k)}$$

La formula precedente, pur essendo una conseguenza immediata della definizione di probabilità condizionata, può risultare utile. Si veda il Probl. 8.

5.1.4. Indipendenza statistica. Se $\mathcal{P}(A|B) = \mathcal{P}(A)$, ciò significa che la conoscenza dell'avverarsi dell'evento B non ha alcuna influenza sulla nostra aspettativa relativamente ad A . Si dice allora che A e B sono statisticamente indipendenti. Esplicitamente la condizione suona così

$$\mathcal{P}(A \cap B) = \mathcal{P}(A) \mathcal{P}(B).$$

La relazione di indipendenza statistica è dunque simmetrica. In generale, dati n eventi A_1, A_2, \dots, A_n , questi si dicono *mutuamente indipendenti* se per ogni combinazione di indici e ogni k , ($2 \leq k \leq n$),

$$\mathcal{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \mathcal{P}(A_{i_1}) \mathcal{P}(A_{i_2}) \dots \mathcal{P}(A_{i_k})$$

Un esempio di eventi che sono indipendenti a due a due ma non *mutuamente* è il seguente: nel lancio di due dadi, $A = \{\text{il primo dado dà un numero dispari}\}$, $B = \{\text{il secondo dado dà un numero dispari}\}$, $C = \{\text{la somma dei dadi è dispari}\}$. Si trova $\mathcal{P}(A) = 1/2$, $\mathcal{P}(B) = 1/2$, $\mathcal{P}(C) = 1/2$. Ora si verifica facilmente che $\mathcal{P}(A \cap B) = 1/4$, $\mathcal{P}(A \cap C) = 1/4$, $\mathcal{P}(B \cap C) = 1/4$; tuttavia $\mathcal{P}(A \cap B \cap C) = 0$! Dunque i tre eventi *non* sono mutuamente indipendenti.

5.2. Variabili aleatorie

Dato uno spazio di probabilità (Ω, \mathcal{P}) si definisce *variabile aleatoria* (random variable) una funzione a valori reali

$$X: \Omega \rightarrow \mathcal{R}$$

La funzione $\mathcal{P}(x' < X < x'')$ è detta *distribuzione di probabilità* della v.a. X . Le variabili aleatorie costituiscono delle coordinate che permettono di individuare gli eventi di Ω in modo simile alle coordinate che si utilizzano in geometria. Ad es., l'evento rappresentato dall'estrazione di un $K\heartsuit$ può essere messo in corrispondenza con una variabile aleatoria a valori interi nell'intervallo $[1-32]$ e il re di cuori corrisponde al valore 7. Avendo valori reali, le variabili aleatorie ammettono tutta una serie di operazioni aritmetiche che hanno un significato probabilistico/statistico. Vediamo alcune di queste operazioni.

Se X è una v.a., si definisce il valore di aspettazione (o valore medio)

$$\mathbb{M}[X] = \sum_{\omega \in \Omega} X(\omega) \mathcal{P}(\omega).$$

Inoltre come è familiare dalla statistica, si definisce la *varianza*

$$\mathbb{D}[X] = \mathbb{M}[(X - \mathbb{M}[X])^2]$$

e la *skewness*

$$\mathbb{S}[X] = \mathbb{M}[(X - \mathbb{M}[X])^3] / \mathbb{D}[X]^{3/2}$$

che dà una misura della asimmetria della v.a. Si definisce poi la *deviazione standard* come

$$\sigma[X] = \sqrt{\mathbb{D}[X]}.$$

5.2.1. Funzione di distribuzione. Per ogni variabile aleatoria si definisce la *funzione di distribuzione* (*f.d.d.*) attraverso la formula

$$\Phi(x) = \mathcal{P}(X < x).$$

La *f.d.d.* è non-decrescente. Vediamo alcuni esempi:

ESEMPIO 5-3 (Bernoulli). La distribuzione di Bernoulli è la più semplice possibile. Lo spazio degli eventi ha solo due elementi ω_1 e ω_2 con $\mathcal{P}(\omega_1) = q, \mathcal{P}(\omega_2) = p$; p e q sono non negativi e a somma uno. Una variabile aleatoria si può definire con $X(\omega_1) = 0, X(\omega_2) = 1$. Più in generale si può considerare lo spazio di probabilità consistente in n bits b_1, b_2, \dots, b_n , statisticamente indipendenti, ognuno dei quali possa assumere il valore 0 o 1 con probabilità q e p , $p + q = 1$; si definiscano le variabili X_i in modo che $X_i(b_1, b_2, \dots, b_n) = b_i$; possiamo pensare alle X_i come il valore dei bit in un banco di memoria a n bit. La funzione di distribuzione di ogni X_i è semplicemente

$$\Phi(x) = \begin{cases} q & 0 \leq x < 1 \\ 1 & x \geq 1. \end{cases}$$



La *funzione di distribuzione congiunta* per due variabili aleatorie si definisce attraverso la relazione

$$\Phi(x, y) = \mathcal{P}(\{X < x\} \cap \{Y < y\})$$

e in generale questa definizione si estende in modo naturale a un numero qualunque di variabili aleatorie. Si vedano altri esempi in Appendice. Nota la *f.d.d.*, il valore di aspettazione è calcolabile attraverso l'integrale

$$\mathbb{M}[X] = \int_{-\infty}^{\infty} x \, d\Phi[x].$$

Questa definizione si applica sia a variabili discrete (come Bernoulli o Poisson) sia a variabili continue (come Gauss). Il significato del differenziale nella formula precedente è da interpretare nel senso delle distribuzioni (più precisamente l'integrale è nel senso di Stieltjes). In molti casi si trova che la *f.d.d.* per una v.a. continua può essere rappresentata attraverso una *densità di probabilità*, ossia

$$\Phi[x] = \int_{-\infty}^x \rho(y) \, dy$$

(è la densità ρ ad es. che è collegata in meccanica ondulatoria alla funzione d'onda ψ attraverso la relazione di Born $\rho = |\psi|^2$).

5.2.2. Funzione generatrice. Sia X una v.a. a valori interi. Consideriamo l'espressione

$$F(z) = \mathbb{M}[z^X] = \sum_i p_i z^i.$$

essendo $p_i = \mathcal{P}(X = i)$. Si tratta di un polinomio nella variabile z che contiene in sé ovviamente tutte le informazioni su X . Ad es. si avrà

$$\mathbb{M}[X] = \sum_i i p_i = F'(1)$$

$$\mathbb{D}[X] = \mathbb{M}[(X - \langle X \rangle)^2] = F''(1) + F'(1) - F'(1)^2.$$

La funzione F viene chiamata *funzione generatrice* della v.a. X . Per variabili a valori continui si utilizza comunemente una definizione un po' differente

$$F(t) = \mathbb{M}[e^{itX}] = \int e^{itx} d\mathcal{P}(x) = \int \rho(x) e^{itx} dx.$$

dalla quale si ottiene immediatamente

$$\mathbb{M}[X^k] = F^{(k)}(0)$$

e cioè lo sviluppo in serie di Taylor della funzione generatrice $F(t)$ ha come coefficienti tutti i momenti della distribuzione (divisi per $k!$).

5.2.3. Variabili indipendenti. Il concetto di *indipendenza statistica* si estende in modo naturale alle variabili aleatorie. Siano X e Y variabili aleatorie definite sullo stesso spazio di probabilità; si dirà che esse sono statisticamente indipendenti se ciò è vero per tutti i sottoinsiemi $X < x$ e $Y < y$. In concreto ciò equivale a dire che la f.d.d. congiunta è fattorizzata:

$$\Phi(x, y) = \Phi(x) \Phi(y)$$

Un indicatore importante per due variabili aleatorie è la loro *correlazione* definita attraverso

$$(5.2) \quad C(X, Y) = \mathbb{M}[XY] - \mathbb{M}[X] \mathbb{M}[Y]$$

Se X e Y sono indipendenti si ha ovviamente $C(X, Y) = 0$.

ESEMPIO 5-4. Due variabili gaussiane sono definite del tutto in generale dalla f.d.d.

$$\Phi(x, y) = \int_{-\infty}^x \int_{-\infty}^y \mathcal{N} \exp\{-\alpha x^2 - \beta y^2 + 2\gamma xy\}.$$

dove la costante \mathcal{N} è fissata in modo che $\Phi \rightarrow 1$ per $x \rightarrow \infty, y \rightarrow \infty$. (Si ha $\mathcal{N} = \sqrt{\alpha\beta - \gamma^2}/\pi$, con la condizione che $\gamma^2 < \alpha\beta$). Si calcoli la covarianza $C(X, Y)$ e si dimostri che essa risulta proporzionale a γ . ◀

5.2.4. Somma di variabili aleatorie. Le v.a. si possono manipolare come le funzioni reali dell'analisi matematica. Molto spesso si trova necessario *sommare* v.a.: ad es. il risultato del lancio di due dadi coincide con la somma di due v.a. D_1 e D_2 : siamo abituati a considerare le facce del dado identificate con il numero corrispondente, ma concettualmente il numero è una variabile aleatoria che identifica la faccia corrispondente, così come la coordinata cartesiana x non è il punto geometrico ma misura la sua distanza da un punto fisso. Ora la somma $D_1 + D_2$ assume valori nell'insieme $2, 3, \dots, 12$. La sua f.d.d. si determina facilmente:

$$\Phi(x) = \begin{cases} 0 & x \leq 2 \\ 1/36 & 2 < x \leq 3 \\ 3/36 & 3 < x \leq 4 \\ \dots & \\ 35/36 & 11 < x \leq 12 \\ 1 & 12 < x \end{cases}$$

In generale la f.d.d. per la somma $X + Y$ è definita da

$$\Phi_{X+Y}(z) = \mathcal{P}(X + Y < z).$$

Per v.a. indipendenti la formula si può calcolare in modo esplicito. Consideriamo il caso discreto, come nel caso del lancio dei dadi. Sia $p(n) = \mathcal{P}(D_i = n) = 1/6$; si ha ovviamente

$$\mathcal{P}(D_1 + D_2 = n) = \sum_{i=1}^6 p(i) \sum_{j=1}^6 p(j) \delta_{i+j=n} = \sum_{i=1}^6 p(i)p(n-i) = (p * p)(n)$$

ossia la funzione di probabilità della somma di v.a. indipendenti è data dalla convoluzione delle rispettive funzioni di probabilità. In modo analogo si trattano le v.a. a valori continui.

Possiamo ora apprezzare l'utilità della definizione di funzione generatrice. Si trova che la f.g. per la somma $D_1 + D_2$ è data da

$$F_{12}(z) = \sum_{i,j} p_i p_j z^{i+j} = F(z)^2;$$

ossia la f.g. per la somma di v.a. indipendenti si ottiene moltiplicando le funzioni generatrici degli addendi. Ciò costituisce una bella facilitazione!

PROBLEMA 5-2. Considerare le v.a. X_i di Bernoulli. Ogni X_i ha una f.g. data da

$$F(z) = q + pz.$$

Calcolare la f.g. per la somma di n variabili di Bernoulli indipendenti, e dedurre il valor medio e la varianza.

SOLUZIONE. Si ha immediatamente $F_n(z) = (q + pz)^n$ e dunque

$$\mathbb{M}[\sum X_i] = F'_n[1] = np$$

$$\mathbb{D}[\sum X_i] = F''_n(1) + F'_n(1) - (F'_n(1))^2 = n(n-1)p^2 + np - (np)^2 = nqp$$

La distribuzione di probabilità della somma $\sum_i X_i$ è la distribuzione binomiale

5.2.5. La legge dei grandi numeri. Sotto questa denominazione vanno numerosi risultati che riguardano valori asintotici di distribuzioni di probabilità quando il numero di gradi di libertà tende all'infinito. L'idea generale è quella della meccanica statistica; un sistema costituito da un numero molto grande di particelle in debole interazione può essere studiato facilmente sfruttando le semplificazioni che si presentano nel limite di infinite particelle. Per variabili aleatorie un caso importante è quello della somma di un grande numero di v.a. indipendenti e tutte con la medesima distribuzione di probabilità. L'esempio più semplice è quello della variabile di Bernoulli $X = \sum X_i$. La sua funzione generatrice è semplicemente $(q + pz)^n$. Il valore medio è dato da np quindi conviene considerare la v.a. centrata $X - \mathbb{M}[X]$, a media zero. Tuttavia non è ancora possibile prendere esplicitamente il limite $n \rightarrow \infty$ in quanto la deviazione standard vale nqp e cresce indefinitivamente. Si può allora considerare la v.a. normalizzata

$$(5.3) \quad \eta = \frac{X - \mathbb{M}[X]}{\sqrt{\mathbb{D}[X]}}$$

che ha perciò media zero e varianza uno. Calcoliamo la sua funzione generatrice nella forma adatta per una distribuzione continua

$$(5.4) \quad F_\eta(x) = \mathbb{M}[e^{it\eta}] = e^{-inpt/\sqrt{nqp}} \mathbb{M}[e^{itX/\sqrt{nqp}}].$$

Adesso però possiamo utilizzare la formula che riduce il calcolo alla potenza n -esima della f.g. per una singola v.a.:

$$(5.5) \quad F_\eta(x) = e^{-i\frac{np}{\sqrt{nqp}}t} \left(q + pe^{i\frac{t}{\sqrt{nqp}}} \right)^n$$

$$(5.6) \quad = \left(q e^{-i\sqrt{\frac{p}{qn}}t} + p e^{-i\sqrt{\frac{q}{pn}}t} \right)^n$$

Osserviamo ora che, nel limite $n \rightarrow \infty$, la funzione F_η ammette un semplice sviluppo asintotico che si ottiene sviluppando le funzioni esponenziali:

$$(5.7) \quad F_\eta = \left(1 - \frac{t^2}{2n} + O(n^{-3/2}) \right)^n$$

e pertanto al limite $n \rightarrow \infty$ la f.g. si riduce a una distribuzione normale di Gauss. Abbiamo ricavato un risultato, noto come legge dei grandi numeri, o teorema di DeMoivre-Laplace, secondo

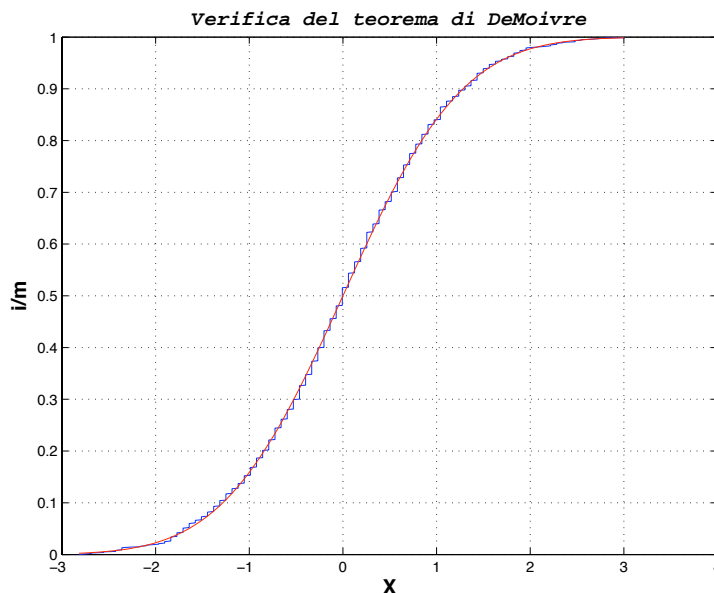


FIGURA 5-3. De Moivre: legge dei grandi numeri, istogramma con $\text{erfc}()$ sovrapposta

cui la somma η definita nell'Eq. (5.3) nel limite di n molto grande è ben descritta da una distribuzione normale. Il risultato è molto generale. Non dipende dalla natura dettagliata della distribuzione di probabilità della singola X_i . Il risultato vale anche, sotto opportune restrizioni, nel caso in cui la distribuzione di X_i possa dipendere da i . Si veda [Roz69]. Il programma `demoivre.m` dimostra attraverso una simulazione lo stesso risultato. Si costruisce una matrice di 0 e 1 distribuiti secondo Bernoulli e la somma sulle colonne corrisponde alla v.a. X . Si analizza l'istogramma e si può controllare di quanto differisce da una distribuzione normale in funzione delle dimensioni dell'esperimento. In Fig.5-3 la linea continua è data dalla funzione $\text{erf}()$, che riproduce la distribuzione Gaussiana (cumulata), cioè $\int_{-\infty}^x \exp\{-\frac{1}{2}y^2\}/\sqrt{2\pi}$. I risultati della simulazione siano $X_1, X_2, \dots, X_i, \dots, X_m$, ordinati in modo crescente, vengono riportati sul diagramma sull'asse delle ascisse e sulle ordinate si riporta l'indice scalato (i/m).

5.3. Processi aleatori

5.3.1. Catene di Markov discrete. Iniziamo a studiare le probabilità in un contesto *dinamico*. Partiamo da un esempio classico, il modello di Ehrenfest. Ci sono due urne che contengono complessivamente N palle, ciascuna identificata da un numero progressivo. Con cadenza regolare si estrae un numero a caso tra uno e N , e la palla che corrisponde al numero estratto passa dall'urna in cui si trova all'altra. Fissata la condizione iniziale (ad es. tutte le palle sono concentrate in una stessa urna), si tratta di determinare la probabilità $P(n, t)$ di trovare nella prima urna n palle all'istante t . Questo modello è rappresentativo di un'intera classe di modelli, denominati *catene di Markov*. L'aspetto essenziale del modello è costituito dal fatto che per una situazione nota all'istante t il futuro non dipende da quello che è successo prima di t . Il processo è determinato da una singola funzione $P(n, t+1|m, t)$ che rappresenta la probabilità di trovare lo stato n al tempo t posto che lo stato sia m all'istante precedente. Questa funzione è detta *probabilità di transizione*. Per il modello di Ehrenfest si trova

$$(5.8) \quad P(n, t+1|m, t) = \begin{cases} 0 & |n - m| > 1 \\ m/N & m = n + 1 \\ 1 - m/N & n = m + 1 \end{cases} .$$

Si può rappresentare la probabilità di transizione con una matrice $N \times N$, $M_{nm} = P(n, t + 1 | m, t)$. Per fissare le idee, se $N = 5$ si trova

$$M = \begin{pmatrix} 0 & 1/5 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2/5 & 0 & 0 & 0 \\ 0 & 4/5 & 0 & 3/5 & 0 & 0 \\ 0 & 0 & 3/5 & 0 & 4/5 & 0 \\ 0 & 0 & 0 & 2/5 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1/5 & 0 \end{pmatrix}$$

Si nota che la matrice è caratterizzata dal fatto che ogni elemento di matrice è non-negativo (deve rappresentare una probabilità) e inoltre la somma degli elementi di matrice su ogni colonna vale esattamente uno. Questo vale in generale: ogni matrice che rappresenta la probabilità di transizione di una catena di Markov deve soddisfare queste condizioni e viene chiamata *matrice stocastica*.

$$(5.9) \quad M_{ij} \geq 0, \quad \sum_i M_{ij} = 1.$$

Se consideriamo la matrice trasposta M^T la condizione (5.9) implica che il vettore con tutte componenti uguali a uno è autovettore con autovalore uno. Dunque $\det(1 - M^T) = 0$, ma il determinante di una matrice è uguale a quello della trasposta, dunque anche M ammette l'autovalore uno. Inoltre si può vedere che tutti gli autovalori di M devono essere localizzati nel disco di raggio uno. Ciò discende dal teorema di Gershgorin [QSS00], riportato in Appendice.

La matrice M permette di calcolare il vettore di probabilità ad ogni tempo t (intero) secondo la semplice formula

$$p(t) = M^t p(0).$$

Ammettiamo che M sia *diagonalizzabile*[§]. Allora vale la decomposizione spettrale

$$M = \sum_{\lambda \in \sigma(M)} \lambda \mathbb{P}_\lambda$$

e quindi

$$M^t = \sum_{\lambda \in \sigma(M)} \lambda^t \mathbb{P}_\lambda.$$

Per tempi molto grandi gli autovalori in modulo inferiori a uno decadono esponenzialmente e sopravvivono solo quelli in modulo uguali a uno. L'autovalore uno esiste sempre; caso per caso possono esistere altri autovalori $e^{i\phi}$. Si avrà allora

$$p(t) \underset{t \rightarrow \infty}{\sim} \mathbb{P}_1 p(0) + \sum_{\phi} e^{i\phi t} \mathbb{P}_{e^{i\phi}} p(0).$$

I termini oscillanti si mediano a zero e quindi il processo tende al limite rappresentato dall'autovettore appartenente all'autovalore uno, che rappresenta la *distribuzione stazionaria*. Sotto condizioni più restrittive gli autovalori di M sono interni al cerchio unitario, ad eccezione del singolo autovalore uno, e quindi la convergenza del processo è assicurata indipendentemente dalla media temporale. Queste condizioni sono enunciate nel teorema di *Perron-Frobenius* riportato in appendice (si veda [Var62, Gan59]).

PROBLEMA 5-3 (Random walk). *Si consideri il random walk in una sola dimensione. Il sistema comporta infiniti stati e quindi l'analisi non può basarsi su quanto detto finora, in particolare i teoremi di Perron e Frobenius non sono applicabili. Dunque è necessario studiare il problema con metodi ad hoc. Si ha $P_{nm} = \frac{1}{2}(\delta_{n,m+1} + \delta_{n,n+1})$. Il calcolo della potenza t -esima*

[§]Quello che è noto in generale sullo spettro di una matrice stocastica irriducibile è che gli autovalori di modulo uno si identificano con le radici dell'unità per un qualche intero r e sono semplici. Per gli autovalori di modulo inferiore a uno non si ha una descrizione così precisa.

di P si risolve con la trasformata di Fourier. Infatti se consideriamo il vettore $\phi(k)_n = \exp\{ink\}$ si verifica immediatamente che

$$P_{nm} \phi(k)_m = \cos k \phi(k)_n$$

Un qualunque vettore u_n si può rappresentare attraverso lo sviluppo di Fourier:

$$u_n = \int_{-\pi}^{\pi} c(k) \phi(k)_n \frac{dk}{2\pi}$$

$$c(k) = \sum_{n=-\infty}^{\infty} \phi(-k)_n u_n$$

e nel caso che si parta dal vettore iniziale $u_n = \delta_{n,j}$, per un dato j , si avrà $c(k) = \phi(-k)_j$. Troviamo allora che l'applicazione di P è data semplicemente da

$$(P^t u)_n = \int_{-\pi}^{\pi} \phi(-k)_j \phi(k)_n (\cos k)^t \frac{dk}{2\pi}$$

e quindi per l'elemento $j \rightarrow j$ vale l'espressione

$$(P^t)_{jj} = \int_{-\pi}^{\pi} (\cos k)^t \frac{dk}{2\pi}.$$

La funzione generatrice $G_j(s)$ (dall'Eq.(5.2)) è quindi

$$(5.10) \quad G_j(s) = \int_{-\pi}^{\pi} \frac{s \cos k}{1 - s \cos k} \frac{dk}{2\pi}$$

Il calcolo offre

$$G_j(s) = \frac{1}{\sqrt{1-s^2}} - 1, \quad F_j(s) = 1 - \sqrt{1-s^2}.$$

Si ottiene allora che i) il camminatore ritorna con probabilità uno al punto di partenza e ii) il tempo medio di ritorno è infinito (dato che la derivata in $s = 1$ diverge), un risultato dovuto a G. Polya. La stessa proprietà vale per il random walk nel piano, mentre in dimensione tre o superiore la probabilità di ritorno risulta inferiore a uno (si veda [ID89]). Si può infatti vedere facilmente che l'integrale (5.10) diventa

$$G_j(s) = \int_{-\pi}^{\pi} \frac{dk_1}{2\pi} \cdots \int_{-\pi}^{\pi} \frac{dk_n}{2\pi} \frac{s \sum \cos k_i}{n - s \sum \cos k_i}$$

e la convergenza o divergenza dell'integrale per k piccolo (in gergo nell'infrarosso) si può decidere anche senza calcolarlo esplicitamente. Ponendo $s = 1$, l'integrando presenta un polo del secondo ordine a $k = 0$. Se $n \geq 3$ l'integrale è finito, il che implica che il camminatore ha solo una probabilità finita di ritornare al punto di partenza ($\int k^{n-1} dk/k^2 < \infty$ per $n \geq 3$).

PROBLEMA 5-4 (Il problema della coda al botteghino). Il problema affrontato nel Probl.4, senza disporre di strumenti specifici, si può ricondurre al problema del random walk. Il numero di biglietti da 5 euro presenti in cassa varia di ± 1 esattamente come nel random walk in una dimensione. Si tratta di valutare la percentuale di cammini che ritornano all'origine senza mai passare per valori negativi. Si è visto che ad ogni cammino di questo tipo corrisponde un cammino con due passi in più tale da non passare mai per lo zero prima dell'ultimo passo. Quindi possiamo contare i cammini che ci interessano attraverso un problema di "primo ritorno" all'origine. Abbiamo visto che la probabilità di ritornare per la prima volta all'origine dopo t passi ha come funzione generatrice $\mathcal{F}(s)$, Eq. (5-3)

$$\mathcal{F}(s) = 1 - \sqrt{1-s^2} = 2 \sum_{n=0}^{\infty} \binom{2n}{n} (s/2)^{2n+2} / (n+1)$$

da cui si deduce che la probabilità di tornare per la prima volta all'origine dopo $2(n+1)$ passi è uguale a $2^{-2n-1} \binom{2n}{n} / (n+1)$. Tenendo conto che il ritorno all'origine può avvenire restando sempre a valori positivi oppure negativi, e che ci sono in totale 2^{2n+2} cammini lunghi $2n+2$, il numero di cammini che si mantengono a valori positivi prima di tornare all'origine è $\binom{2n}{n} / (n+1)$ e questo coincide, per quanto visto nel Probl. 5-4, al numero di cammini lunghi $2n$ che non scendono mai sotto lo zero. Si noterà che ci sono diversi modi di impostare il problema della coda: se sappiamo a priori che il numero di biglietti da 5 e da 10 euro sono uguali, allora la probabilità è da conteggiare limitatamente ai cammini che comunque partono e arrivano a zero in $2n$ passi; in questo caso la probabilità di “non-attesa” è appunto $1/(n+1)$. Se invece non abbiamo alcuna informazione sulla distribuzione dei biglietti il numero di cammini possibili è dato da 2^{2n} . Dimostrare che in questo caso la probabilità di non essere costretti ad attendere, per mancanza di resto, diminuisce di un ulteriore fattore $2^{-2n} \binom{2n}{n} \approx 1/\sqrt{\pi n} (1 - \frac{1}{8n} + \frac{1}{128n^2} + O(n^{-3}))$.

PROBLEMA 5-5. Un problema di campionamento che si può presentare spesso (pensiamo a sistemi di controllo di qualità nei prodotti di qualunque tipo) è il seguente (da [Knu98], § 3.4.2): supponiamo di avere un insieme di N oggetti da cui si vuole estrarre un campione di $n \ll N$ per controllo. Gli N oggetti però non sono disponibili subito allo stesso momento ma si rendono disponibili nel corso del tempo (nei prossimi 30 giorni, ad es.), e tuttavia non è il caso di attendere ad iniziare i controlli di qualità. Allora il problema è quello di selezionare il campione senza introdurre bias di alcun genere. La prima idea, cioè selezionare l' m -esimo oggetto con probabilità n/N si rivela poco soddisfacente: non c'è la sicurezza di avere esattamente n campioni, e la fluttuazione intorno al valore centrale può essere non trascurabile (vale precisamente $\sqrt{n(1-n/N)}$). Provare a trovare la soluzione e verificarla con una simulazione, prima di continuare.

SOLUZIONE. La probabilità di selezione dipende da quanti campioni si sono già estratti: se dopo $0 \leq k < N$ estrazioni si sono già selezionati c campioni con $0 \leq c \leq n$, allora all'estrazione successiva si accetta l'esemplare con probabilità

$$P(n, N, c, k) = \frac{n - c}{N - k}.$$

L'idea è del tutto intuitiva: dopo k estrazioni e c esemplari prescelti si comincia di nuovo con $N \leftarrow N - k$ e $n \leftarrow n - c$. Questo garantisce che alla fine il numero di esemplari è precisamente n ; resta però da dimostrare che in questo modo ogni campione viene selezionato davvero a caso (per esempio non si accumulano i campioni verso la fine della selezione). Il programma `matlab` aiuta a decidere positivamente:

```

function X = campioni(n,N)
% Selezionare a caso n campioni da una
% popolazione di N elementi
X=[];
nc=n;                % campioni ancora da estrarre
Nc=N;                % esemplari residui
while nc>0,
    if(rand < nc/Nc)
        X=[X, Nc]; % ok selezionato un campione
        nc = nc-1;
    end
    Nc=Nc-1;        % rimangono Nc esemplari
end

```

5.3.2. Catene di Markov continue. Studiamo ora un'altra varietà di processi aleatori. Sempre limitandoci a sistemi con un numero finito di stati possibili, esistono casi in cui la probabilità $p_n(t)$ dipende da un parametro continuo t . Si pensi ad esempio a un insieme di nuclidi radioattivi A_1, A_2, \dots, A_N di cui si conosce il ritmo di decadimento $\lambda_{i \rightarrow j}$; la densità $p_n(t)$ soddisfa l'equazione differenziale

$$\frac{dp_n}{dt} = - \sum_{m \neq n} \lambda_{n \rightarrow m} p_n(t) + \sum_{m \neq n} \lambda_{m \rightarrow n} p_m(t)$$

che risulta della forma generale

$$\frac{dp_n}{dt} = \sum_m H_{n,m} p_m(t).$$

Affinché questa equazione rappresenti un processo di Markov che conservi la probabilità totale uguale a uno, la matrice H dovrà soddisfare un vincolo:

$$0 = \sum_n \frac{dp_n}{dt} = \sum_{n,m} H_{n,m} p_m(t).$$

che è automaticamente soddisfatto se $\sum_n H_{n,m} = 0$, cioè la somma degli elementi di ogni colonna somma a zero. Si richiederà inoltre che gli elementi fuori diagonale $H_{n,m}$ siano positivi (rappresentano una probabilità di transizione nell'unità di tempo), mentre i termini diagonali devono saldare il bilancio della colonna e quindi sono negativi.

5.3.3. Proprietà di convergenza. Si tratta di controllare le proprietà spettrali della matrice H ; infatti la soluzione $p(t) = \exp(tH)p(0)$ tenderà a un *unico* limite $p(\infty)$ se sono soddisfatte due condizioni: *i)* esiste un unico autovettore corrispondente all'autovalore zero e *ii)* gli altri autovalori hanno parte reale negativa. Ora, la condizione che la somma sulle colonne dia zero implica immediatamente che lo zero appartiene allo spettro (il vettore $(1, 1, \dots, 1)$ è autovettore sinistro), ma in generale esisteranno più autovettori, a meno che H non soddisfi l'ulteriore proprietà di essere *irriducibile* (in modo analogo alle matrici stocastiche del Cap. 5.3.1). Per la seconda condizione, si può concludere del tutto in generale grazie al teorema di Gershgorin (vedi 5.5). In base a questo teorema si ha infatti che lo spettro è contenuto nell'unione dei cerchi (si ricordi che $H_{ii} \leq 0$)

$$\mathcal{C}_i = \{z \in \mathbb{C} \mid |z - H_{ii}| \leq \sum_{j \neq i} H_{ji} = -H_{ii}\}$$

Dunque ognuno dei cerchi \mathcal{C}_i ha centro sull'asse reale negativo e passa per l'origine. In conclusione, detto \bar{h} il massimo valore assoluto degli elementi sulla diagonale di H , lo spettro è compreso nel cerchio di raggio \bar{h} centrato in $-\bar{h}$. Se consideriamo l'evoluzione infinitesimale

$$p(t + \delta t) \sim p(t) + \delta t H p(t)$$

questa è equivalente a una catena di Markov con matrice di transizione

$$M = 1 + \delta t H$$

a patto di scegliere un passo δt sufficientemente piccolo, e precisamente $\delta t \leq 1/\bar{h}$. Il problema è quindi ricondotto a quello di catene discrete. Nel caso irriducibile, la distribuzione asintotica è perciò sempre data dalla soluzione di $H = 0$.

Si noti che, viceversa, ogni catena di Markov discreta si può trasformare in un processo continuo: per ogni salto della catena si introduca un certo numero n_h (halt) di pause in cui il processo si arresta; la decisione di arresto sia presa con probabilità $n_h/(n_h + 1)$, ne segue allora un processo governato da

$$p_{t+1} = (1 - r)p_t + r M p_t$$

avendo introdotto il “parametro di rilassamento” $r = 1/(n_h + 1)$. È chiaro che se facciamo crescere n_h l’equazione tenderà a rappresentare un’evoluzione continua governata dall’equazione

$$\frac{dp_t}{dt} = (M - 1)p_t$$

Si vede che la matrice $M - 1$ è proprio una matrice a elementi positivi fuori diagonale e soddisfa la condizione che la somma su ogni colonna si annulla. È quindi del tipo considerato nel paragrafo precedente.

5.3.4. Il metodo MonteCarlo. Con il termine MonteCarlo si intende una vasta categoria di tecniche e algoritmi che permettono di simulare numericamente un processo fisico (totalmente o parzialmente) governato dal caso. L’interesse può risiedere nel processo stesso (ad es. si vuole studiare la dinamica dei neutroni in un reattore) oppure il processo in sé è fittizio, non rappresenta cioè una situazione fisica reale, ma il risultato dell’analisi statistica coincide con una qualche grandezza collegata, questa sì, alla fisica. Ad es. lo studio di un sistema di meccanica statistica in *quattro* dimensioni, pur non corrispondendo a un sistema fisico reale, permette di calcolare grandezze che hanno un’interpretazione in termini di masse o ampiezze di transizione in fisica quantistica. Le tecniche MonteCarlo nascono con i primi calcolatori elettronici, un nome che ricorre molto di frequente (Metropolis) è quello di un collaboratore di Enrico Fermi e John Von Neumann (si veda a proposito [Mal03]). Si deve in gran parte a Mark Kac l’intuizione secondo cui lo studio di processi stocastici attraverso la simulazione numerica può fornire metodi di calcolo per la meccanica quantistica.

Un problema tipico che si presenta è quello di generare successioni di stati individuati da coordinate x_1, x_2, \dots, x_n e che sono distribuiti a caso secondo una probabilità

$$(5.11) \quad P(x \in \mathcal{S}) = \frac{1}{Z} \int_{x \in \mathcal{S}} e^{-\beta H[x]}.$$

Qui $Z(\beta)$, detta la *funzione di partizione*, assicura la normalizzazione della distribuzione. Questo problema è tipico per la meccanica statistica (classica). La generazione di una successione di stati scelti in modo indipendente permette di valutare le varie medie statistiche (energia media, magnetizzazione, etc.). Lo stesso schema è applicato in fisica quantistica (si veda [Reb83]). Data l’importanza di disporre di sistemi di calcolo affidabili che permettano di risolvere problemi di questo genere (ad es. in teoria della *interazioni forti*) si sono sviluppate varie tecniche che vanno sotto il nome di *Metropolis*, *Heat-bath*, *Langevin*, *Hybrid-MC*. Questi argomenti interessano più che altro lo specialista. In queste lezioni ci limitiamo a illustrare alcuni aspetti della problematica. Si consiglia di consultare le lezioni di A. Pelissetto [Pel94] e di Sokal [Sok97] per approfondire il tema.

5.3.5. Come realizzare successioni casuali. Il problema ridotto all’osso è il seguente. Data una lista di eventi $\Omega = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N\}$ con probabilità p_1, p_2, \dots, p_N , come si fa a generare una successione di eventi che siano *mutuamente indipendenti* e distribuiti secondo la probabilità $\mathcal{P}(\varepsilon_i) = p_i$?

L’ingrediente essenziale per ogni algoritmo di simulazione è costituito da un generatore di numeri (pseudo)-random distribuiti uniformemente nell’intervallo $[0, 1)$. Ammetteremo di disporre di questo ingrediente, presente in qualunque libreria matematica, con l’avvertenza che non tutti i generatori hanno le stesse qualità. Non esistono in realtà generatori *perfetti*, che soddisfino cioè alla condizione di uniformità e indipendenza statistica. Ogni generatore (ad es. `rand()`, `drand48()` in C o analoghi in Fortran, `rand()` in Matlab, etc.) produce una successione $r_1, r_2, \dots, r_n, \dots$ che innanzitutto è data da numeri in virgola mobile a precisione finita (usualmente 64 bit ossia 15 cifre) che quindi non possono riprodurre se non un sottoinsieme dei reali. Ma la questione più grave è che, inevitabilmente, i numeri generati sono *correlati*. Molta cura

si è posta nel disegno di algoritmi che diminuiscano al di sotto di limiti accettabili questa correlazione. Si veda [Knu98], Vol.II. Ammettiamo dunque di avere un buon generatore $rand()$, con le precauzioni del caso.[¶] Si tratta ora di trasformare la successione

$$r_1, r_2, \dots, r_n, \dots \rightarrow \varepsilon_{i_1}, \varepsilon_{i_2}, \dots, \varepsilon_{i_k}, \dots$$

in modo da realizzare un campionamento (una simulazione) di (Ω, \mathcal{P}) . In genere serviranno *molte* estrazioni di numeri casuali per costruire un esemplare in Ω (si pensi ad es. a una distribuzione di punti in \mathcal{R}^d , serviranno allora almeno d chiamate a $rand()$).

La prima idea è quella (poco efficiente) di organizzare una cascata di diramazioni del tipo

```

    ???
    r = rand;
    if r < p_1
        return 1
    else if r < p_1 + p_2
        return 2
    else if r < p_1 + p_2 + p_3
        return 3
    .....
    end if

```

A parte la goffaggine dello schema, appare evidente che il metodo risulta poco efficiente. Sarà necessario in media effettuare $\mathcal{O}(N)$ controlli logici per estrarre un singolo risultato. Una formulazione più compatta è la seguente, ma è solo un miglioramento estetico:

```

    ?
    q=cumsum(p)-p(N);    % q=[0,p(1),p(1)+p(2),...,1-p(N)]
    x=[];
    for j=1:Nsample
        r=rand;
        x=[x; max(find(q<r))];
    end

```

Il tempo di esecuzione per $N_{\text{sample}} = 10000, N = 10$ è dell'ordine di 1/2 sec. Descriviamo ora un algoritmo che permette di estrarre un esemplare a caso con un numero di operazioni $\mathcal{O}(1)$, il che abbassa drasticamente il tempo di esecuzione!

PROBLEMA 5-6 (Il problema dei portalettere). *All'ufficio postale sono in giacenza L lettere indirizzate a N diversi destinatari. Il Direttore assume N persone per risolvere rapidamente il problema. A complicare le cose succede, com'è naturale, che le lettere non sono equamente distribuite: ℓ_1 sono indirizzate a \mathcal{I}_1 , ℓ_2 a \mathcal{I}_2 , etc. Tuttavia per accordi sindacali ogni portalettere deve al massimo visitare due indirizzi e la quantità di lettere deve essere rigorosamente la stessa per tutti. Come fa il direttore ad assegnare le buste rispettando i vincoli sindacali?*

🔗 Il problema si trova su [Knu98] (alias method).

SOLUZIONE. *Il metodo consiste nel suddividere ricorsivamente le buste secondo un principio semplice: organizziamo gli indirizzi in ordine crescente di volume di posta, cioè ammettiamo che $\ell_1 \leq \ell_2 \leq \dots \leq \ell_N$. Assegniamo ℓ_1 buste a \mathbb{P}_1 . Se fosse $\ell_1 = \ell$ saremmo nel caso banale, in cui tutti gli indirizzi hanno lo stesso numero di buste - ma abbiamo escluso questo caso in quanto non pone problema di sorta. Dunque abbiamo $\ell_1 < \ell$. Allora \mathbb{P}_1 può prendersi anche $\ell - \ell_1$ buste, e la scelta giusta è quella di prenderle con l'indirizzo \mathcal{I}_N , quello con il numero maggiore di buste. A questo punto \mathbb{P}_1 è al completo e può partire. Restano $N - 1$ portalettere e $(N - 1)\ell$ buste recanti*

[¶]Generatori di ottime qualità statistiche sono stati costruiti in tempi recenti (Lüscher, Mersenne Twister, etc.). **matlab** implementa questi nuovi algoritmi.

$(N - 1)$ indirizzi (perchè il primo indirizzo è ormai completamente soddisfatto). Ma allora, una volta rimessi in ordine gli indirizzi secondo ℓ_k crescenti ci ritroviamo al problema iniziale con N diminuito di uno. Si applica la stessa strategia fino ad esaurire gli indirizzi (e i portalettere). ◀

È evidente che l'idea si può tradurre in un metodo per estrarre efficientemente un valore $i \in \{1, 2, \dots, N\}$ con probabilità p_1, p_2, \dots, p_N . Il problema si mappa su quello dei portalettere con l'equivalenza $p_i = \ell_i/L$. Se al portalettere \mathbb{P}_j spettano ℓ'_j lettere per \mathcal{I}_j si pone $\pi_j = \ell'_j N/L$. Ad ogni portalettere viene assegnato un *secondo indirizzo* \mathcal{I}'_j . Avendo perciò a disposizione le tabelle $\{\pi_j, \mathcal{I}'_j \mid j = 1 : N\}$ si può definire così un metodo per estrarre a caso dall'insieme $\{\mathcal{I}_j\}$ con probabilità $\{p_j\}$:

- i) si estrae un numero intero j a caso tra 1 e N
- ii) si estrae un numero reale r tra 0 e 1
- iii) se $r < \pi_j$ si sceglie \mathcal{I}_j
- iv) altrimenti si sceglie \mathcal{I}'_j (l'altro indirizzo)

Il metodo è noto come *il metodo degli alias*. Si veda [Knu98], vol.II, oppure [Pel94]. L'implementazione `matlab` si trova nel package accluso (`KnuthAliasMethod`). Si può utilizzare anche per il caso di distribuzioni continue; a questo scopo si suddivide il campo di definizione in tanti canali (*bins*) in modo che la scelta del canale viene effettuata con il metodo degli alias e all'interno del canale si approssima la distribuzione con una funzione lineare per la quale esiste un algoritmo veloce. Il lavoro addizionale costituito dalla costruzione delle tabelle è ampiamente ripagato dalla velocità dell'algoritmo (a patto che si lavori in assetto costante, cioè che la distribuzione di probabilità sia sempre la stessa!).

5.3.6. MonteCarlo dinamico. Per le applicazioni recenti del MonteCarlo interessanti la Fisica Teorica, il problema centrale consiste nel generare una successione casuale di configurazioni di fase per un sistema di meccanica statistica distribuite secondo la formula di Gibbs. In sostanza si utilizza una catena di Markov irriducibile e, sperabilmente, con buone proprietà di convergenza, per campionare i punti $x \in \mathcal{S}$ nello spazio delle fasi \mathcal{S} e determinare così una stima dell'integrale di fase Eq.(5.11). Un algoritmo MC specifico consiste in una catena di Markov con probabilità di transizione $P_{x \rightarrow x'}$, costruita in modo tale che la distribuzione di equilibrio $\pi(x') = \sum_x \pi(x) P_{x \rightarrow x'}$ sia data da $\pi(x) = \exp\{-\beta H[x]\}/Z(\beta)$. Concentriamo l'attenzione su un particolare problema di meccanica statistica, quello di sistemi di spin distribuiti su un reticolo spaziale. Un tipico esempio è il modello di Heisenberg: in ogni punto di un reticolo $x \in \mathbb{Z}^3$ è definito un vettore unitario $\mathbf{n}(x)$ che rappresenta la direzione del momento magnetico $\boldsymbol{\mu} = \mu \mathbf{n}$. L'energia di interazione tra due dipoli magnetici decresce molto velocemente con la distanza, quindi in una prima schematizzazione si considera che ogni magnete elementare interagisca solo con i magneti più vicini (i *primi vicini*) e l'energia totale assume perciò la forma

$$H(\mathbf{n}) = -J \sum_{\langle xy \rangle} \boldsymbol{\mu}(x) \cdot \boldsymbol{\mu}(y)$$

essendo J una costante e la notazione $\langle xy \rangle$ viene a significare che la somma va estesa alle coppie di primi vicini. In presenza di un campo magnetico esterno \mathbf{B} all'energia va aggiunto anche un termine $-\sum_x \mathbf{B} \cdot \boldsymbol{\mu}(x)$. Ora il valore di una grandezza fisica, quale la *magnetizzazione*, è esprimibile attraverso un integrale del tipo

$$\langle \mathcal{M}(x) \rangle = Z(\beta)^{-1} \prod_{x \in \mathbb{Z}^3} \int d\boldsymbol{\mu}(x) e^{-\beta H(\mathbf{n})} \boldsymbol{\mu}(x).$$

In generale si desidera valutare il valor medio di una qualunque grandezza fisica, funzione definita nello spazio delle fasi,

$$\langle \mathcal{O}(\mu) \rangle = Z(\beta)^{-1} \prod_{x \in \mathbb{Z}^3} \int d\mu(x) e^{-\beta H(\mu)} \mathcal{O}(\mu(x)) .$$

In casi speciali, per modelli più semplici, si è riusciti a trovare metodi per calcolare integrali così complessi per via analitica. Tuttavia, nella gran parte dei modelli interessanti, la via della simulazione MonteCarlo rappresenta l'unica risorsa.

Sono stati concepiti vari metodi per realizzare questo obiettivo. Forse il più semplice concettualmente quello del cosiddetto “heat bath”. Nella sua forma più semplice il metodo consiste nel visitare sequenzialmente tutti i punti del reticolo (in gergo *una sweep* e di volta in volta determinare $\mu(x)$ come se fosse l'unica variabile dinamica (tutte le altre variabili si tengono temporaneamente congelate). Dato che $\mu(x)$ interagisce solo con un numero limitato di primi vicini, lo spazio di probabilità è costituito dai punti della sfera con densità di probabilità proporzionale a

$$\exp\{-\beta J \mu(x) \cdot \sum_{k=1}^6 \mu(x_k) - \mathbf{B} \cdot \mu(x)\}$$

dove x_k indica i primi vicini di x .

A titolo di esempio, prendiamo un modello più semplice che basterà a chiarire l'idea. Il modello di **Ising** prevede in ogni punto di un reticolo spaziale una variabile a due soli valori ± 1 , essenzialmente una restrizione del modello di Heisenberg in cui si vincoli ogni magnete ad avere la stessa direzione in modo che sopravviva come variabile solo l'orientazione. Se consideriamo di più il caso di un reticolo piano il modello si riduce al seguente

$$(5.12) \quad \langle \mathcal{O}(\mu) \rangle = Z(\beta)^{-1} \prod_{x \in \mathbb{Z}^2} \sum_{\mu(x)=\pm 1} e^{\beta J \sum_{\langle xy \rangle} \mu(x)\mu(y)} \mathcal{O}(\mu(x)) .$$

Per un qualunque spin $\mu(x)$ i primi vicini sono quattro e per una data configurazione si possono riscontrare le situazioni seguenti

$$\begin{array}{cccccccccccc} \begin{array}{c} + \\ \square \\ + \end{array} & \begin{array}{c} + \\ \square \\ - \end{array} & \begin{array}{c} - \\ \square \\ + \end{array} & \begin{array}{c} - \\ \square \\ - \end{array} & \begin{array}{c} + \\ \square \\ + \end{array} & \begin{array}{c} - \\ \square \\ - \end{array} & \begin{array}{c} + \\ \square \\ - \end{array} & \begin{array}{c} - \\ \square \\ + \end{array} & \begin{array}{c} - \\ \square \\ - \end{array} & \begin{array}{c} + \\ \square \\ + \end{array} & \dots \end{array}$$

più tutte le altre simili per un totale di 2^4 . La probabilità di $\mu(x)$ dipende solo dalla somma degli spin primi vicini; detta $s(x) = J \beta \sum_y \mu(y) \delta_{\|x-y\|,1}$ la probabilità sarà data semplicemente da

$$\mathcal{P}(\mu(x)=\pm 1) = \frac{e^{\pm s(x)}}{e^{s(x)} + e^{-s(x)}} .$$

La somma $s(x)$ può assumere solo cinque valori differenti, ed è quindi molto facile e rapido assegnare a $\mu(x)$ un valore in accordo con questa formula. Si applica questa ricetta a tutti gli spin successivamente e poi si ricomincia. Ovviamente è necessario limitare il sistema a un numero finito di punti individuando una regione finita del piano. In gergo questo si dice “fissare il cut-off infrarosso”. Non è consigliabile effettuare la sweep punto per punto, ma conviene agire su un sottoreticolo di variabili disaccoppiate, ad esempio tutte quelle di posto pari (somma delle coordinate) e in seguito quelle di posto dispari. Il problema si presta a essere affrontato *in parallelo*. Un grande numero di sweeps sarà necessario per generare una configurazione del modello non troppo correlata alla iniziale. Infatti certe configurazioni locali tendono a mantenersi stabili: per β grande gli spin tendono a favorire valori uguali tra primi vicini (spin allineati) ed esiste perciò un tempo caratteristico per ogni modello che è necessario attendere prima che la correlazione sia diminuita di un fattore $1/e$ - si considera cioè che la correlazione avrà una dipendenza dal “tempo” della catena di Markov del tipo $C(t, t') = \exp(-|t-t'|/\tau)$. Tanto più piccolo è τ , il *tempo di autocorrelazione*, tanto più efficiente sarà il processo di generazione di configurazioni statisticamente indipendenti. Molto studio è stato messo nella analisi e nello sviluppo di algoritmi nel tentativo di abbattere il tempo di autocorrelazione. È

un buon investimento, perchè un algoritmo che migliori l'efficienza rappresenta un risparmio di tempo e/o di denaro. Per chi sia interessato a sviluppi recenti si consiglia di consultare i *Proceedings* delle conferenze di Teorie di Gauge su Reticolo, che escono con cadenza annuale, in particolare [Sok91].

5.3.7. Equazione di Langevin. Con il lavoro di Parisi e Wu [PW81] i metodi delle equazioni differenziali stocastiche (EDS) entrano a far parte della cassetta degli attrezzi del fisico interessato alle interazioni forti. In precedenza le EDS erano assunte a base dell'approccio di Nelson [Nel67] alla Meccanica Quantistica. La tecnica delle EDS si era tuttavia sviluppata già nella prima metà del XX secolo, con applicazioni a vari problemi di Fisica Classica. Qui daremo solo un accenno, rimandando a [Wax54, DH87].

L'idea nella sua formulazione più semplice è quella di studiare un'equazione differenziale del tipo

$$(5.13) \quad d\mathbf{x}(t) = \mathbf{b}(\mathbf{x}(t))dt + \sigma(\mathbf{x}(t))dw(t)$$

dove $\mathbf{x}(t)$, e $\mathbf{b}(\mathbf{x}(t))$ sono in generale funzioni a valori in \mathcal{R}^n , $\sigma(\mathbf{x}(t))$ è una funzione reale positiva e $w(t)$ è il *moto browniano standard* in n dimensioni. La definizione di quest'ultimo può richiedere una lunga preparazione matematica (vedi [Nel67]), ma per quanto ci interessa si può definire per via puramente assiomatica:

(a) per ogni scelta di istanti t_1, t_2, \dots, t_N si ha che $\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_N)$ sono variabili mutuamente Gaussiane che soddisfano le seguenti proprietà:

(b) $\mathbb{M}[\mathbf{x}(t)] = 0$

(c) $\mathbb{M}[\mathbf{x}_i(t) \mathbf{x}_j(s)] = \delta_{ij} \min(t, s)$.

Ne segue immediatamente che gli incrementi $\mathbf{x}(t_1) - \mathbf{x}(t_2)$ e $\mathbf{x}(t_3) - \mathbf{x}(t_4)$ sono statisticamente indipendenti se gli intervalli (t_1, t_2) e (t_3, t_4) non si sovrappongono, mentre

$$(5.14) \quad \mathbb{M}[|\mathbf{x}(t+h) - \mathbf{x}(t)|^2] = h$$

Vedremo ora brevemente come si può risolvere in pratica l'equazione Eq. (5.13). Prima però osserviamo che, secondo un teorema di cui non daremo dimostrazione, il processo $\mathbf{x}(t)$ è caratterizzato da una densità di probabilità in funzione del tempo che soddisfa l'equazione di Chapman-Kolmogorov [PR69]:

$$(5.15) \quad \frac{\partial P(\mathbf{x}, t)}{\partial t} = \frac{1}{2} \Delta (\sigma(\mathbf{x})^2 P(\mathbf{x}, t)) - \nabla \cdot (\mathbf{b}(\mathbf{x}) P(\mathbf{x}, t))$$

Ora, c'è una classe particolare di possibili $\mathbf{b}(\mathbf{x})$ (in gergo il "drift") che permette di ricavare la soluzione asintotica per t grande: è il caso in cui $\mathbf{b}(\mathbf{x})$ è un campo irrotazionale e nel contempo σ è una costante. Si ha infatti che, posto $\mathbf{b}(\mathbf{x}) = \nabla \rho(\mathbf{x})$, il membro di destra dell'equazione diventa

$$\begin{aligned} & \frac{1}{2} \Delta (\sigma^2 P(\mathbf{x}, t)) - \nabla \cdot (\nabla \rho(\mathbf{x}) P(\mathbf{x}, t)) \\ &= \nabla \cdot \left(\frac{1}{2} \sigma^2 \nabla P(\mathbf{x}, t) - P(\mathbf{x}, t) \nabla \rho \right) \end{aligned}$$

e quindi si annulla per $P = \exp\{-2\rho(\mathbf{x})/\sigma^2\}$. Ma allora è possibile costruire una EDS con soluzione stazionaria preassegnata scegliendo il drift secondo la ricetta $\mathbf{b} = \nabla \rho, \rho = \sigma^2 \log(P)/2$. Il caso che interessa in meccanica statistica e in teoria quantistica è quello in cui

$$P(\mathbf{x}) = Z^{-1} \exp\{-\beta H(\mathbf{x})\}$$

e quindi

$$\mathbf{b}(\mathbf{x}) = -\nabla H(\mathbf{x}), \quad \sigma^2 = 2/\beta.$$

Dunque è possibile costruire un processo aleatorio $x(t)$ che è garantito convergere per tempi grandi a una distribuzione stazionaria che può essere scelta a piacere. Questo significa che l'EDS può essere utilizzata come algoritmo di simulazione MonteCarlo.

5.3.8. Equazione di Langevin e Meccanica Quantistica. Vi sono almeno due contesti in cui le EDS sono utili nello studio della Meccanica Quantistica. Nelson [Nel67] ha mostrato che l'equazione di Schroedinger è matematicamente equivalente ad una famiglia di processi diffusivi, governati da EDS con $\sigma \propto \sqrt{\hbar}$ e $b(x, t) = \nabla \log(\psi(x, t))$. Da un punto di vista pratico, questo fatto non comporta grossi benefici, in quanto anche solo per scrivere l'EDS è necessario conoscere la soluzione dell'Equazione di Schroedinger o un suo sistema equivalente di equazioni non lineari. L'altro punto di vista è invece quello della “quantizzazione stocastica” nel senso di Parisi-Wu. Il nocciolo della questione è dato dal fatto che l'EDS con drift $b(x) = \nabla H(x)$ permette di valutare medie statistiche rispetto alla misura di Gibbs $\exp(-\beta H(x))$, e inoltre la Meccanica Quantistica formulata a tempo immaginario coincide con la Meccanica Statistica classica. Per l'equazione di Schroedinger di una particella scalare in un grado di libertà e energia potenziale $V(x)$ la funzione $H(x)$ si scrive

$$H(x) = \int_{-\infty}^{\infty} \left(\frac{1}{2} \dot{x}^2 + V(x(t)) \right) dt$$

e l'equazione di Langevin assume la forma

$$\frac{\partial x(t, \tau)}{\partial \tau} = \frac{\partial^2 x(t, \tau)}{\partial t^2} - V'(x(t)) + \sigma \eta(t, \tau)$$

Il fatto importante è che *si possono ottenere informazioni sullo spettro di energia e su elementi di matrice quantistici dallo studio di funzioni di correlazione* del processo aleatorio. Ad es. si ha (vedi App. 5.9)

$$(5.16) \quad \mathbb{M}[x(t)x(s)] \sim \sum_{E \in \sigma} |\langle E_0 | q | E \rangle|^2 e^{-(E-E_0)|t-s|/\hbar}$$

dove q è l'operatore di posizione, $|E_0\rangle$ rappresenta lo stato fondamentale e la somma è estesa a tutto lo spettro di energia σ . Questa corrispondenza offre la possibilità di stimare le caratteristiche spettrali di sistemi quantistici complessi per i quali un calcolo attraverso l'equazione di Schroedinger non sarebbe possibile. Ovviamente l'approccio MonteCarlo introduce necessariamente degli errori (statistici e sistematici) che devono essere valutati e ridotti al minimo. Ma questo riguarda un'altra serie di lezioni.

Appendice IV

5.4. Distribuzioni di probabilità elementari

- *Distribuzione di Bernoulli: vedi distribuzione binomiale con $n = 1$*
- *Distribuzione binomiale*

Rappresenta la distribuzione di probabilità per la somma di n variabili aleatorie indipendenti a valori 0 e 1 ciascuna con probabilità $\mathcal{P}(0) = p$.

Spazio degli eventi elementari: $\Omega = \{0, 1, 2, \dots, n\}$

Probabilità: $\mathcal{P}(k) = \binom{n}{k} q^{n-k} p^k$, $(p + q = 1)$

v.a.: $X(i) = i$

Funzione generatrice: $F(z) = (q + pz)^n$

Media: $\mathbb{M}[X] = np$

Dev.St.: $\mathbb{D}[X] = npq$

Skewness: $\mathbb{S}[X] = \frac{1-2p}{\sqrt{npq}}$

- *Distribuzione di Poisson (Si ottiene dalla binomiale nel limite $np \rightarrow \lambda$, $n \rightarrow \infty$)*

Spazio degli eventi elementari: $\Omega : \mathbb{Z}$

Probabilità: $\mathcal{P}(k) = e^{-\lambda} \lambda^k / k!$, $\lambda > 0$

Funzione generatrice: $F(z) = e^{\lambda(z-1)}$

Media: $\mathbb{M}[x] = \lambda$

Dev.St.: $\mathbb{D}[x] = \lambda$

Skewness: $1/\sqrt{\lambda}$

- *Distribuzione di Gauss*

Spazio degli eventi elementari: $\Omega = \mathbb{R}$

Densità di probabilità: $\rho(x) = \frac{e^{-\frac{1}{2}x^2/\sigma}}{\sqrt{2\pi\sigma}}$, $\sigma > 0$

Funzione generatrice: $F(t) = \int_{-\infty}^{\infty} dx e^{itx} \rho(x) = e^{-\frac{1}{2}\sigma t^2}$

v.a.: $X(x) = x$

Media: $\mathbb{M}[X] = 0$

Dev.St.: $\mathbb{D}[X] = \sigma$

Skewness: 0

- *Gauss multidimensionale*

Spazio degli eventi elementari: $\Omega = \mathbb{R}^n$

Densità di probabilità: $\rho(\mathbf{x}) = \frac{e^{-\frac{1}{2} \langle \mathbf{x} | \mathbf{A}^{-1} | \mathbf{x} \rangle}}{(2\pi)^{n/2} \sqrt{\det \mathbf{A}}}$, $\mathbf{A} \gg 0$

Funzione generatrice: $F(\mathbf{t}) = \int_{-\infty}^{\infty} d\mathbf{x} e^{i \langle \mathbf{t} | \mathbf{x} \rangle} \rho(\mathbf{x}) = e^{-\frac{1}{2} \langle \mathbf{t} | \mathbf{A} | \mathbf{t} \rangle}$

v.a.: $X_i(\mathbf{x}) = x_i$

Media: $\mathbb{M}[X_i] = 0$

Dev.St.: $\langle X_i X_j \rangle = A_{ij}$.

La matrice \mathbf{A} di dimensione n è *positiva definita*, ossia $\langle \mathbf{x} | \mathbf{A} | \mathbf{x} \rangle > 0$ per ogni vettore \mathbf{x} non nullo. Il calcolo della funzione di distribuzione per una densità di probabilità gaussiana in \mathbb{R}^n si effettua come segue $\mathbf{M} \equiv \mathbf{A}^{-1} = \mathbf{R}^{-1} \mathbf{\Lambda} \mathbf{R}$, $\mathbf{\Lambda}$ diagonale:

$$\begin{aligned} F(\xi) &= \mathcal{N} \int e^{i \langle \xi | \mathbf{x} \rangle - \frac{1}{2} \langle \mathbf{x} | \mathbf{M} \mathbf{x} \rangle} [d\mathbf{x}] = \\ &= \mathcal{N} \int e^{i \langle \xi | \mathbf{R}^{-1} \mathbf{x}' \rangle - \frac{1}{2} \langle \mathbf{R}^{-1} \mathbf{x}' | \mathbf{M} \mathbf{R}^{-1} \mathbf{x}' \rangle} [d\mathbf{x}'] = \\ &= \mathcal{N} \int e^{i \langle \xi | \mathbf{R}^{-1} \mathbf{x}' \rangle - \frac{1}{2} \langle \mathbf{x}' | \mathbf{R} \mathbf{M} \mathbf{R}^{-1} \mathbf{x}' \rangle} [d\mathbf{x}'] = \\ &= \mathcal{N} \int e^{i \langle \mathbf{R} \xi | \mathbf{x}' \rangle - \frac{1}{2} \langle \mathbf{x}' | \mathbf{\Lambda} \mathbf{x}' \rangle} [d\mathbf{x}'] = \\ &= \mathcal{N} \prod_{j=1}^n \int e^{i (\mathbf{R} \xi)_j x_j - \frac{1}{2} \Lambda_{jj} x_j^2} dx_j = \\ &= \mathcal{N} \prod_{j=1}^n \left(\sqrt{\frac{2\pi}{\Lambda_{jj}}} e^{-\frac{1}{2\Lambda_{jj}} (\mathbf{R} \xi)_j^2} \right) = \\ &= \mathcal{N} \frac{(2\pi)^{\frac{n}{2}}}{\sqrt{\det \mathbf{M}}} e^{-\frac{1}{2} \langle \mathbf{R} \xi | \mathbf{\Lambda}^{-1} (\mathbf{R} \xi) \rangle} = \mathcal{N} \frac{(2\pi)^{\frac{n}{2}}}{\sqrt{\det \mathbf{M}}} e^{-\frac{1}{2} \langle \xi | \mathbf{M}^{-1} \xi \rangle} \end{aligned}$$

e imponendo $F(0) = 1$ si trova

$$\begin{aligned} F(\xi) &= e^{-\frac{1}{2} \langle \xi | \mathbf{M}^{-1} \xi \rangle} \\ \rho(\mathbf{x}) &= \frac{\sqrt{\det \mathbf{M}}}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2} \langle \mathbf{x} | \mathbf{M} \mathbf{x} \rangle} \end{aligned}$$

5.5. I teoremi di Gershgorin e Frobenius

Per i due teoremi che seguono si veda [Var62].

Teor. 5.5.1 TEOREMA 5.5.1 (Gershgorin). *Sia \mathbf{M} una matrice qualunque $N \times N$. Lo spettro della matrice è contenuto nell'unione dei cerchi*

$$\{z \in \mathbb{C} \mid |z - M_{ii}| \leq \sum_{j \neq i} |M_{ji}|\}.$$

Corollario: se la somma $\sum_j M_{ji} = 1$, si trova $|z - M_{ii}| \leq 1 - M_{ii}$. Tutti i “cerchi di Gershgorin” sono allora contenuti nel cerchio unitario e passano tutti per il punto $z = 1$. Il teorema seguente, fondamentale nell'analisi delle catene di Markov finite, si trova sul testo di Varga oppure sul trattato di Gantmacher ([Gan59], Vol.II, Ch. XIII).

TEOREMA 5.5.2 (Perron–Frobenius). *Si assuma che la matrice \mathbf{A} di dimensione N a elementi non-negativi sia irriducibile, cioè non esiste alcuna permutazione degli indici che la trasformi nella forma ridotta*

Teor. 5.5.2

$$\begin{pmatrix} \mathbf{A}_{r \times r} & \mathbf{A}_{r \times s} \\ 0 & \mathbf{A}_{s \times s} \end{pmatrix}$$

con $r + s = N$. Allora valgono le seguenti proprietà:

- lo spettro di \mathbf{A} contiene un autovalore reale positivo uguale al raggio spettrale $\rho(\mathbf{A})$;
- in corrispondenza di questo autovalore positivo esiste un unico autovettore avente tutte componenti positive;
- il raggio spettrale $\rho(\mathbf{A})$ è una funzione strettamente crescente degli elementi di matrice;
- se k è il numero di autovalori con modulo pari al raggio spettrale la matrice è detta ciclica di indice k e può essere ridotta attraverso una permutazione degli indici alla forma

$$\begin{pmatrix} 0 & \mathbf{B}_{12} & 0 & \dots & 0 \\ 0 & 0 & \mathbf{B}_{23} & \dots & 0 \\ \vdots & \vdots & \ddots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{B}_{r-1,r} \\ \mathbf{B}_{r,1} & 0 & 0 & \dots & 0 \end{pmatrix}.$$

- Gli autovalori di modulo $\rho(\mathbf{A})$, per una matrice ciclica di indice k , sono semplici e coincidono con le radici k -esime dell'unità $\exp\{2\pi i j/r\}$, e in generale per ogni autovalore λ anche $\lambda \exp\{2\pi i j/r\}$ appartiene allo spettro.

Il teorema determina le proprietà asintotiche del processo di Markov. Infatti per tempi lunghi la matrice di transizione sarà data da

$$P^t \underset{t \rightarrow \infty}{\sim} \sum_{j=0}^{r-1} e^{2\pi i j t/r} \mathbb{P}_j + \mathcal{O}(\lambda_{\text{sub}}^t),$$

dove i termini trascurabili sono associati a tutti gli autovalori in modulo inferiori a uno e λ_{sub} è quello di modulo più grande. Si afferma di solito che il limite della distribuzione di probabilità esiste solo per il caso *primitivo*, cioè per $r = 1$. Questo è a rigore vero, ma irrilevante per le applicazioni dove quello che conta è il comportamento del processo *mediato su un tempo finito*. Infatti è chiaro che il contributo alla probabilità a tempi lunghi degli autovettori con autovalore $\exp\{2\pi i j/r\}$, $j > 0$ risulta fortemente attenuato se procediamo ad una media su un tempo finito (ricordiamo che per una qualunque radice dell'unità $\omega \neq 1$ si ha $\sum_{k=1}^r \omega^k = 0$). Si veda anche il seguente § 5.8.

5.6. Distribuzione di Polya

Partiamo dalla distribuzione di Bernoulli: in un contenitore sono presenti w palle bianche e b nere, $b + w = N$. Si compiono n estrazioni a caso, riponendo ogni volta la palla estratta nell'urna, e si registra il numero di palle bianche sulle n estrazioni totali. La probabilità si trova con il metodo del Cap.5.2.3. Una variante possibile è quella di Polya: nello stesso contenitore si estrae a caso una palla e, una volta rimessa al suo posto, le si aggiungono s palle dello stesso colore ($s \geq 0$). Questa procedura è ripetuta n volte. Sia $X_{n,s}$ la v.a. che prende il valore k se in queste n estrazioni vengono selezionate k palle bianche. Il caso $s = 0$ coincide con la distribuzione binomiale con $p = w/N$. Un semplice programma `matlab` permette di esplorare le proprietà della distribuzione:

```

Polya.m
function Wtot=polya(Nsw,n,s,w1,b1)
% Usage: [Wtot,P]=polya(Nsw,n,s,w1,b1)
% ..... omessi dettagli .....
Wtot=[];
for j=1:Nsw      % start Nsw draws from Polya
    N=w1+b1; p=w1/N; W=0; w=w1; b=b1; % initialize
    for k=1:n
        if(rand<p), W=W+1; w=w+s;
        else,      b=b+s;
        end
        N=w+b; p=w/N; % update
    end
    Wtot=[Wtot,W]; % store result
end

```

In Fig. 5-4 un risultato per $n = 100, s = 1, w = 4, b = 3$ con $\mathbb{M}[X_{100,1}] = 57.14$. La distribuzione di Polya è data da

$$P[X_{n,s} = k] = \binom{n}{k} \frac{(w/s)_k (b/s)_{n-k}}{((w+b)/s)_n}$$

dove $(a)_n = \prod_{j=0}^{n-1} (a+j)$ è il simbolo di Pochhammer. Il caso limite $s = 0$ si riduce alla distribuzione di Bernoulli $P[X_{n,0} = k] = \binom{n}{k} (w/(w+b))^k (b/(w+b))^{n-k}$ (si veda [Fis63]).

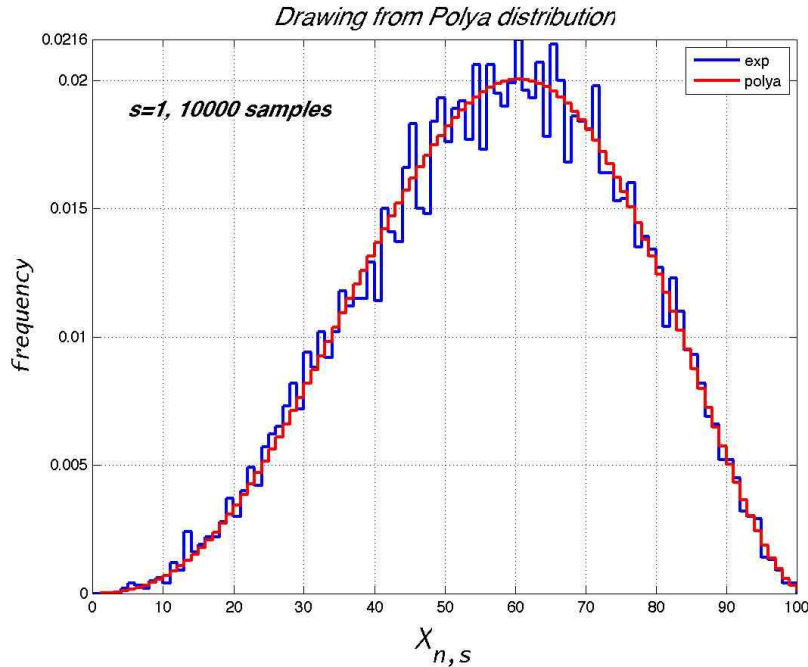


FIGURA 5-4. Distribuzione di Polya

5.7. Soluzione esatta del modello di Ehrenfest

Se $p(n,t)$ è la probabilità che la prima urna contenga n particelle al tempo t , conviene introdurre una funzione generatrice definita come segue:

$$G(x,t) = \sum_{n=0}^N p(n,t) x^n.$$

La relazione di ricorrenza definita dalla probabilità di transizione (5.8)

$$p(n, t+1) = \frac{n+1}{N} p(n+1, t) + \frac{N-n+1}{N} p(n-1, t)$$

(dove il primo coefficiente è da considerare uguale a zero per $n = N$ e il secondo per $n = 0$) si traduce in una relazione per $G(n, t)$:

$$\begin{aligned} G(x, t+1) &= \sum_{n=0}^{N-1} \left(\frac{n+1}{N} p(n+1, t) x^n + \frac{N-n+1}{N} p(n-1, t) x^n \right) \\ &= \sum_{n=0}^{N-1} \left(\frac{1}{N} p(n+1, t) \frac{d}{dx} x^{n+1} + p(n-1, t) x^n - \frac{1}{N} x^2 p(n-1, t) \frac{d}{dx} x^{n-1} \right) \\ &= \frac{1-x^2}{N} \frac{d}{dx} p(x) + x p(x) \end{aligned}$$

La distribuzione di probabilità stazionaria deve soddisfare l'equazione $G(x, t+1) = G(x, t)$ ossia $(1+x) \frac{dG}{dx}(x, t) = N G(x, t)$ che ha come unica soluzione $G(x, t) = \mathcal{N} (1+x)^2$. La normalizzazione ($\sum p(n) = G(1, t) = 1$), impone $\mathcal{N} = 2^{-N}$. Ne segue che la distribuzione di equilibrio è data da $p(n) = 2^{-N} \binom{N}{n}$. Per N molto grande i valori più grandi di probabilità si hanno per $n \approx N/2$. Per esercizio si calcoli $\mathbb{M}[n]$ e $\mathbb{D}[n]$. Con un calcolo analogo si trovano tutti gli altri autovalori della matrice di transizione (precisamente $\{j/N | j = -N, -N+1, \dots, N-1, 1\}$). La presenza dell'autovalore -1 implica che la matrice M di transizione è *ciclica di indice 2*.

5.8. Calcolo del tempo medio di ritorno

Calcoliamo le funzioni generatrici $F_j(s)$ e $G_j(s)$ utilizzando la decomposizione spettrale della matrice P (si veda Pag. 193). Per semplicità ammettiamo che P sia diagonalizzabile, anche se questo non sarebbe strettamente necessario. Sia perciò

$$P = \sum_{\lambda} \lambda \mathbb{P}_{\lambda}$$

dove la somma è estesa a tutto lo spettro di P . Si ha allora

$$G_j(s) = \sum_{\lambda} \frac{s\lambda}{1-s\lambda} (\mathbb{P}_{\lambda})_{jj}$$

Dalla relazione (5.4) si ricava

$$\frac{dF_j}{ds} = \frac{G'_j(s)}{(1+G_j(s))^2}$$

e inserendo la rappresentazione spettrale si ottiene così

$$(5.17) \quad \mu_j = \frac{dF_j}{ds} = \frac{\sum_{\lambda} \frac{\lambda}{(1-s\lambda)^2} (\mathbb{P}_{\lambda})_{jj}}{(\sum_{\lambda} \frac{1}{1-s\lambda} (\mathbb{P}_{\lambda})_{jj})^2}$$

$$(5.18) \quad = \frac{\frac{1}{(1-s)^2} (\mathbb{P}_1)_{jj} + \sum'_{\lambda} \frac{\lambda}{(1-s\lambda)^2} (\mathbb{P}_{\lambda})_{jj}}{(\frac{1}{1-s} (\mathbb{P}_1)_{jj} + \sum'_{\lambda} \frac{1}{1-s\lambda} (\mathbb{P}_{\lambda})_{jj})^2}$$

$$(5.19) \quad = (\mathbb{P}_1)_{jj}^{-1} + \mathcal{O}(1-s) .$$

dove \sum' indica la somma estesa agli autovalori diversi da uno. Ora il proiettore \mathbb{P}_1 è dato da $(\mathbb{P}_1)_{nm} = \mathbf{u}_n^R \mathbf{u}_m^L$ essendo \mathbf{u}^R e \mathbf{u}^L gli autovettori *destri* e *sinistri* di P con autovalore uno. Dato che l'autovettore sinistro ha tutte componenti uguali a uno (dalla la condizione di stocasticità Eq. 5.9) si ha per il limite $s \rightarrow 1$

$$\mu_j = 1/u_j^R$$

Si noti che nel caso di matrice di transizione *irriducibile* tutte le componenti della distribuzione stazionaria u_R sono strettamente positive.

Il codice Mathematica che segue calcola le funzioni generatrici F e G per il modello di Ehrenfest.

```

EhrenfestP[n_] :=
  SparseArray[{{i_, j_}/; i-j==1->1-(j-1)/n, {i_, j_}/; i-j==1->
    i/n}, {n+1, n+1}]
P[n_] := P[n] = Normal[EhrenfestP[n]];
R[n_, s_] := R[n, s] =
  Simplify[s P[n] . Inverse[IdentityMatrix[n+1] - s P[n],
    Method->CofactorExpansion]];
G[n_, s_] := G[n, s] = Simplify[Tr[R[n, s], List]];
F[n_, s_] := F[n, s] = Simplify[G[n, s] / (1 + G[n, s])];
K=15;
Simplify[F[K, s]] /. s -> 1
{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1}
(* ritorno c.p.1 per tutti i casi iniz. *)
Series[D[F[K, s], s], {s, 1, 1}] (* Output omissso *)
mu=Limit[D[H[K, s], s], s->1]
{32768, 32768/15, 32768/105, 32768/455, 32768/1365, 32768/3003,
  32768/5005, 32768/6435, 32768/6435, 32768/5005, 32768/3003,
  32768/1365, 32768/455, 32768/105, 32768/15, 32768}
(* 2^15/Binomial(15, k) *)

```

I valori di μ_n coincidono con i reciproci delle componenti della distribuzione stazionaria (Sec. 5.7). Si ha qui un esempio di processo che sebbene tenda all'equilibrio prevede che ogni stato si ripresenti con probabilità uno ma su tempi medi che possono diventare molto grandi (immaginare cosa sia un tempo 2^{1023} anche ammettendo che il passo temporale sia 10^{-10} sec!), con il che, su scale temporali limitate, il processo appare irreversibile.

5.9. Dimostrazione della Eq. (5.16)

Sia $q(t)$ l'operatore di posizione in descrizione di Heisenberg, cioè $q(t) = \exp(i\mathcal{H}t) q \exp(-i\mathcal{H}t)$. Passiamo a tempo immaginario e calcoliamo

$$\begin{aligned}
 \langle E_0 | q(it) q(is) | E_0 \rangle &= \langle E_0 | q e^{-\mathcal{H}(s-t)} q | E_0 \rangle e^{E_0(s-t)} \\
 &= \sum_n \langle E_0 | q | E_n \rangle \langle E_n | q | E_0 \rangle e^{-(E_n - E_0)(s-t)} \\
 &= \sum_n |\langle E_0 | q | E_n \rangle|^2 e^{-(E_n - E_0)(s-t)}
 \end{aligned}$$

D'altra parte, scelto un vettore arbitrario $|\psi\rangle$, utilizzando la relazione

$$|E_0\rangle = \lim_{T \rightarrow \infty} e^{-(\mathcal{H} - E_0)T} |\psi\rangle / \langle E_0 | \psi \rangle$$

si ha anche

$$\begin{aligned}
 \langle E_0 | q(it) q(is) | E_0 \rangle &= \lim_{T \rightarrow \infty} \frac{\langle \psi | e^{-\mathcal{H}T} q e^{-\mathcal{H}(s-t)} q e^{-\mathcal{H}T} | \psi \rangle}{\langle \psi | e^{-\mathcal{H}T} e^{-\mathcal{H}(s-t)} e^{-\mathcal{H}T} | \psi \rangle} \\
 &\sim \frac{\int \mathcal{D}x(.) e^{-\int_{-\infty}^{\infty} d\tau \left(\frac{1}{2} \dot{x}^2 + V(x(\tau)) \right)} x(t)x(s)}{\int \mathcal{D}x(.) e^{-\int_{-\infty}^{\infty} d\tau \left(\frac{1}{2} \dot{x}^2 + V(x(\tau)) \right)}}
 \end{aligned}$$

L'ultimo passaggio presume familiarità con il formalismo dei *path integrals* di Feynman come sviluppati ad es. su [Fey72]. Si veda [Pol77] per l'applicazione al calcolo del gap di energia e per l'estensione a sistemi più interessanti.

5.10. Processi diffusivi

Sia $\mathbf{x}(t) : [0, +\infty) \rightarrow \mathbb{R}^n$ un processo stocastico che evolve nel tempo secondo la seguente equazione:

$$(5.20) \quad d\mathbf{x}(t) = \mathbf{b}(\mathbf{x}(t), t) dt + \sqrt{\sigma(\mathbf{x}(t), t)} d\mathbf{w}(t)$$

dove $\mathbf{b}(\mathbf{x}(t), t)$ è un campo vettoriale, detto campo di *drift*, che agisce in modo deterministico sul sistema; $d\mathbf{w}(t)$ è una variabile casuale distribuita in modo gaussiano con $\mathbb{M}[d\mathbf{w}] = 0$ e

$$\mathbb{M}[d\mathbf{w}(t) d\mathbf{w}(s)] = dt \delta(t - s),$$

La funzione σ rappresenta il coefficiente di diffusione, cosicché $\sqrt{\sigma} d\mathbf{w}(t)$ fornisce il contributo aleatorio allo spostamento infinitesimo, come descritto in figura.

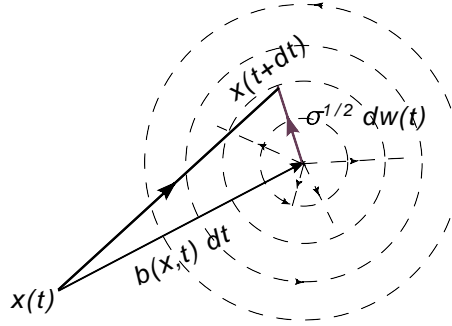


FIGURA 5-5. Lo spostamento $\mathbf{x}(t) \rightarrow \mathbf{x}(t+dt)$ decomposto in parte casuale e in trascinamento (*drift*)

La probabilità che $\mathbf{x}(t)$ si trovi al tempo t in una certa porzione di spazio $\mathcal{I} \subset \mathbb{R}^n$ è data da:

$$\mathcal{P}_{\mathcal{I}}(t) = \int_{\mathcal{I}} \rho(\mathbf{x}, t) d\mathbf{x}$$

dove $\rho(\mathbf{x}, t)$ è la densità di probabilità. Si trova, sotto opportune condizioni, che $\rho(\mathbf{x}, t)$ soddisfa

$$(5.21) \quad \frac{\partial \rho}{\partial t} = \frac{\sigma}{2} \Delta \rho - \nabla \cdot (\rho \mathbf{b})$$

detta *equazione di Kolmogorov*. Il primo addendo rappresenta la parte detta *diffusiva* del processo, mentre il secondo è il classico termine dell'equazione di continuità del campo \mathbf{b} di cui la (5.21) rappresenta una generalizzazione.

Nel caso in cui il drift è il gradiente di una funzione, allora è immediato determinare la distribuzione invariante, cioè la soluzione stazionaria dell'Eq.(5.21). Sia infatti

$$\mathbf{b}(\mathbf{x}) = -\frac{\sigma}{2} \nabla S(\mathbf{x}).$$

Segue che il membro di destra dell'equazione si annulla per $\rho = \exp\{-S\}$. Nell'applicazione alla meccanica quantistica, la funzione S rappresenta l'azione (euclidea, cioè a tempo immaginario) e per consistenza è presente un fattore $1/\hbar$

$$\rho(\mathbf{x}) = \exp\left\{-\frac{1}{\hbar} S(\mathbf{x})\right\}$$

e l'equazione stocastica assume la forma

$$d\mathbf{x}(t) = -\frac{\sigma}{2\mathbf{h}} \nabla S(\mathbf{x}) dt + \sqrt{\sigma} d\mathbf{w}(t)$$

La costante di diffusione σ regola l'evoluzione temporale ma non influenza la soluzione stazionaria, quindi si può scegliere arbitrariamente, ad es. $\sigma = 2\mathbf{h}$, il che semplifica formalmente l'equazione; sono però a disposizione altre scelte che possono rendere più veloce la convergenza alla soluzione stazionaria. Ad esempio è anche possibile scegliere σ uguale ad una matrice (positiva definita); nel caso in cui S sia quadratica $\frac{1}{2}\langle \mathbf{x}, A \mathbf{x} \rangle$, è evidente il vantaggio, come ha suggerito Parisi, di porre $\sigma = A^{-1}$; infatti si può dare il caso in cui gli autovalori di A siano molto diversi come ordine di grandezza, cioè l'indice di condizionamento dato dal rapporto tra l'autovalore più grande e quello più piccolo sia molto grande. In questa situazione il calcolo numerico della soluzione dell'equazione stocastica risulta molto penalizzata (per raggiungere l'equilibrio è necessario attendere i modi più lenti). La scelta di Parisi mette tutti i modi al passo con un notevole vantaggio di efficienza dell'algoritmo ^{||}.

Come esempio affrontiamo la risoluzione di un problema particolare

$$d\mathbf{x}(t) = -A\mathbf{x}dt + \sqrt{2} d\mathbf{w}(t),$$

(dove abbiamo posto per semplicità $\mathbf{h} = 1$), la cui soluzione è

$$(5.22) \quad \mathbf{x}(t) = e^{-At}\mathbf{x}(0) + \int_0^t e^{A(t'-t)} d\mathbf{w}(t')$$

dove $\mathbf{x}(0)$ è la condizione iniziale. Dato che \mathbf{x} risulta sovrapposizione lineare di variabili gaussiane (sempre che lo sia $\mathbf{x}(0)$), la soluzione è data ad ogni istante da una variabile aleatoria gaussiana; per caratterizzarla è allora sufficiente conoscerne la media e la (co-)varianza. Dalla (5.22) si trova $\mathbb{M}[\mathbf{x}(t)] = e^{-At}\mathbb{M}[\mathbf{x}(0)]$, essendo $\mathbb{M}[d\mathbf{w}(t)] = 0$. Essendo A definita positiva si ha $\lim_{t \rightarrow \infty} \mathbb{M}[\mathbf{x}(t)] = 0$, avendo semplicità posto $\mathbf{x}(0) = 0$, il che non influenza il risultato a tempi grandi, e troviamo

$$\begin{aligned} \mathbb{M}[x_i(t) x_j(t)] &= \mathbb{M}\left[\int_0^t (e^{A(t'-t)})_{ik} d\mathbf{w}_k(t') \int_0^t (e^{A(t''-t)})_{jh} d\mathbf{w}_h(t'')\right] \\ &= \int_0^t \int_0^t (e^{A(t'-t)})_{ik} (e^{A(t''-t)})_{jh} \mathbb{M}[d\mathbf{w}_k(t') d\mathbf{w}_h(t'')] \\ &= \int_0^t \int_0^t (e^{A(t'-t)})_{ik} (e^{A(t''-t)})_{jh} 2\delta_{kh} \delta(t' - t'') dt' \\ &= 2 \int_0^t (e^{2A(t'-t)})_{ij} dt' \\ &= \left(\frac{1 - e^{-2At}}{A}\right)_{ij} \end{aligned}$$

Si è sfruttato il fatto che la matrice A è simmetrica. Nel limite per $t \rightarrow \infty$ si trova

$$\mathbb{M}[x_i(t) x_j(t)] \rightarrow (A^{-1})_{ij}$$

Se adottiamo la scelta di Parisi, $\sigma = A^{-1}$, la soluzione diviene

$$\mathbf{x}(t) = \mathbf{x}(0)e^{-t} + \int_0^t e^{t'-t} A^{-1/2} d\mathbf{w}(t')$$

da cui segue

$$\mathbb{M}[x_i(t) x_j(t)] = (A^{-1})_{ij} (1 - e^{-2t})$$

^{||} Il calcolo della radice quadrata $A^{-1/2}$ non pone eccessivi problemi; inoltre è effettuata una volta sola nel corso del calcolo.

Quindi la soluzione dipende dalla scelta di σ , ma non la distribuzione stazionaria, che è data da

$$\rho(x, t) \rightarrow \mathcal{N} \exp\{-\frac{1}{2} \langle x | A | x \rangle\}$$

Un problema interessante è quello di realizzare algoritmi efficienti per studiare numericamente la soluzione di equazioni differenziali stocastiche; teniamo conto infatti che i casi in cui si sa risolvere esattamente l'equazione si riducono a poca cosa (caso lineare o riconducibile a). Un metodo molto semplice è basato sull'osservazione seguente: per integrazione si ha dall'Eq. (5.13) (avendo assunto $\sigma = 1$)

$$x(t + \tau) = x(t) + \int_t^{t+\tau} b(x(t')) dt' + w(t + \tau) - w(t),$$


e dunque la parte casuale è data *esattamente* da un contributo gaussiano con media 0 e varianza $\sqrt{\tau}$. L'integrale invece bisogna approssimarlo in modo accurato. La formula del trapezio è la prima possibilità

$$\int_t^{t+\tau} b(x(t')) dt' \approx \frac{1}{2} \tau (b(x(t)) + b(x(t + \tau))) .$$

Questo tuttavia obbliga a risolvere un'equazione implicita per ottenere $x(t + \tau)$. In alternativa si può approssimare ulteriormente la formula inserendo $x(t + \tau) \approx x(t) + b(x(t))\tau + w(t + \tau) - w(t)$ nella formula del trapezio.

Per altri algoritmi, più accurati, si può consultare [KP99]. Per l'applicazione delle equazioni stocastiche al Monte Carlo si veda invece [DH87].

Problemi

 Il simboli \star e $\star\star$ indicano gli esercizi più impegnativi. Il simbolo $\#$ segnala che la soluzione, almeno parziale, è riportata in fondo al capitolo.

PROB 1. $\#$ Si tirano tre dadi. Si determini la probabilità che la somma dia 11 oppure 12. (Nel '600, un certo DeMère, che passava molto tempo al tavolo da gioco, si era stupito del fatto che $\mathcal{P}(11) > \mathcal{P}(12)$, almeno stando alle sue osservazioni, ma Pascal diede la dimostrazione che il fatto era del tutto prevedibile sulla base di un argomento matematico).

PROB 2. Un cubo di legno ha le facce colorate. Il cubo viene sezionato in dieci parti parallelamente alla faccia $x-y$, poi ancora i dieci parti secondo $x-z$ e infine secondo $y-z$. Qual'è la probabilità che scegliendo a caso uno dei mille cubetti questo abbia due facce colorate?

PROB 3. $\#$ Si sceglie a caso un numero intero compreso tra 1 e 1000. Qual'è la probabilità che il cubo del numero abbia le ultime due cifre uguali a 1? (cioè $n^3 = \star\star 11$).

PROB 4. \star Duecento spettatori si presentano in fila ordinata alla biglietteria del cinema Astra, in possesso per la metà di biglietti da 5 Euro (il costo esatto del biglietto) e l'altra metà di biglietti da 10. La cassa del teatro è inizialmente vuota. Qual'è la probabilità che nessuno nella coda debba aspettare per avere il resto?

PROB 5. $\#$ Una variante dell'esercizio precedente: supponiamo che nella coda ci siano $n+m$ spettatori muniti di banconote da 5 Euro e $n-m$ con 10 Euro. Come varia la probabilità che la coda non incontri interruzioni?

PROB 6. In un gioco a premi, il presentatore mette sul tavolo tre buste di cui una sola contiene il premio. Una volta scelta la busta, al concorrente viene indicata quale tra le due buste rimanenti è sicuramente vuota e gli si offre la possibilità di modificare la scelta. Il concorrente argomenta che a questo punto la probabilità di vincere è al 50% e tanto vale mantenere la scelta iniziale. Tuttavia il concorrente sbaglia a non approfittare dell'offerta. Perché?

PROB 7. $\#$ Il classico problema "della rovina del giocatore": giocando a testa-o-croce un giocatore gioca ogni volta un Euro. Assumendo che il suo capitale iniziale sia di x Euro, il giocatore si ripropone di continuare ad oltranza fino a raggiungere un capitale di m Euro ($m > x$) oppure perde tutto. Qual'è la probabilità di uscire con il portafoglio vuoto? (vedi `gambler.m`).

PROB 8. $\#$ In dieci contenitori si hanno nell'ordine 2 palle bianche e 2 nere nei primi nove, mentre nell'ultimo sono contenute 5 bianche e una nera. Si estrae a caso una palla senza sapere da quale contenitore. Se si è scelta una palla bianca, qual'è la probabilità che ce ne siano altre quattro nello stesso contenitore?

PROB 9. Dimostrare la formula

$$\mathcal{P}(A \cap B \cap C) = \mathcal{P}(A) \mathcal{P}(B|A) \mathcal{P}(C|A \cap B) .$$

PROB 10. Due punti vengono scelti a caso nell'intervallo $(0, L)$. Qual'è la probabilità che la loro distanza non superi ℓ ?

PROB 11. Due urne contengono inizialmente N_A particelle di tipo A e N_B di tipo B ciascuna. Viene spostata a caso una particella dalla prima alla seconda urna e successivamente dalla seconda alla prima. Il processo si ripete indefinitamente. Come varia la probabilità di avere n_A particelle di tipo A nella prima urna (e n_B del secondo tipo)?

PROB 12. Passeggiata aleatoria (random walk—textbf): consideriamo un reticolo nel piano costituito da tutti i punti con coordinate cartesiane intere, $\mathbf{x} = (n_1, n_2)$. Ad ogni istante una particella puntiforme salta a caso dalla sua posizione ad una vicina con uguale probabilità in ognuna delle quattro direzioni. Determinare la probabilità $\mathcal{P}_t(n_1, n_2)$ di trovare la particella nel punto (n_1, n_2) all'istante t .

PROB 13. Sia dato un grafo rappresentato in Fig. 5-6. Si considera un cammino casuale sul grafo che inizia in O . Ad ogni diramazione il cammino sceglie un ramo con la stessa probabilità (tranne che ritornare indietro). Qual'è la probabilità di raggiungere il nodo A ?

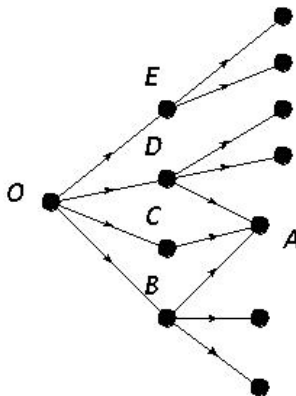


FIGURA 5-6. Il grafo del Probl. 13

PROB 14. [#]**Ritorno al punto di partenza:** consideriamo una catena di Markov su N stati, caratterizzato dalla probabilità di transizione P_{ij} . Si suppone che tutti gli stati possano essere raggiunti prima o poi, cioè che P sia irriducibile*. Ci si chiede con quale probabilità partendo dallo stato j si tornerà a j dopo un numero qualunque di passi. Inoltre ha interesse sapere quanto tempo in media deve passare perchè ciò avvenga.

PROB 15 (si veda [Gin00]). Si studi il caso di un sistema a tre stati A, B, C la cui funzione di probabilità soddisfa l'equazione di evoluzione (α, β e γ sono costanti positive)

$$\begin{cases} dp_A/dt = -\alpha p_A + \gamma p_C \\ dp_B/dt = \alpha p_A - \beta p_B \\ dp_C/dt = \beta p_B - \gamma p_C \end{cases}$$

*Questo significa in altre parole che nello spazio degli stati del processo non esistono sottoinsiemi da cui il processo non possa uscire.

e se ne trovi, se esiste, la soluzione asintotica per tempi grandi. Se non si riesce a trovare la soluzione, si esplori il problema per qualche valore particolare dei parametri attraverso una simulazione numerica (e proprio come ultima risorsa, prima di esaminare la soluzione, si veda il codice `aug_problem.m`).

PROB 16 (Problema di Ehrenfest continuo). Si prenda il problema delle due urne del Cap. 5.3.1, con una variante: l'estrazione a caso non avviene con regolarità ogni τ secondi, piuttosto l'estrazione avviene con probabilità λdt nell'intervallo dt . Siamo dunque nel caso prefigurato, con un parametro di rilassamento

$$p(t + dt) = (1 - \lambda dt) p(t) + \lambda dt M_E p(t)$$

dove abbiamo indicato con M_E la matrice stocastica del modello discreto delle due urne. Si ha dunque un processo di Markov in tempo continuo governato dalla matrice $H_E = M_E - 1$. Dal risultato dell'App. 5.7 si può verificare che lo spettro di H_E è contenuto nel cerchio di raggio uno centrato in -1 . Per visualizzare l'evoluzione del vettore di probabilità si può utilizzare il codice `aug_problem3.m`.

PROB 17. Dimostrare che la matrice di transizione del modello di Ehrenfest soddisfa l'equazione $M T + T M = 0$ essendo T la matrice diagonale con elementi di matrice $T_{n,n} = (-1)^n$. (ciò implica anche che lo spettro di autovalori di M deve essere simmetrico intorno all'origine delle coordinate).

PROB 18. Dalla relazione Eq. 5.17 ricavare i primi due termini dello sviluppo di Taylor per $\mu_j(s)$ e dedurne la varianza del tempo medio di ritorno $\mathcal{D}(\tau_j)$.

PROB 19. #* Si consideri una matrice M $n \times n$ Hermitiana i cui elementi di matrice sono estratti a caso da una distribuzione Gaussiana normale. Determinare la densità di probabilità $P(\lambda \in [\chi, \chi + d\chi])$ per gli autovalori, nel limite di n molto grande. La distribuzione è nota come "legge del semicerchio di Wigner" ([Wig67, Meh89]).

PROB 20. #** Costruire un algoritmo che permetta di estrarre matrici unitarie a caso, uniformemente rispetto alla misura invariante $[dU] = [d, g U] = [d U g]$.

PROB 21. #* Costruire un algoritmo che permetta di estrarre matrici stocastiche a caso, secondo la convenzione: gli elementi di ciascuna colonna formano un vettore casuale, con componenti non-negative, distribuito uniformemente nella regione $\sum_i M_{ij} = 1$.

PROB 22. Una pulce esegue un salto ogni minuto primo da uno all'altro di quattro cani C_1, \dots, C_4 . La decisione avviene (apparentemente) a caso con probabilità di salto $C_j \rightarrow C_k$ data dalla matrice stocastica

$$P(C_j \rightarrow C_k) = \begin{pmatrix} 0 & 1/4 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 1/4 & 1/2 & 1/2 \end{pmatrix}$$

Si calcoli la probabilità che la pulce si trovi sul cane C_1 dopo un'ora dalla prima rilevazione che la trovava su C_2 . Dimostrare che il risultato è praticamente indipendente dal cane di partenza.

PROB 23.[#] Si consideri il gruppo S_n delle permutazioni su n oggetti. Si chiede qual'è la probabilità che una permutazione scelta a caso non lasci fisso alcun elemento? (Una variante del problema dice: 40 signori lasciano il proprio cappello all'ingresso del teatro. All'uscita viene a mancare l'energia elettrica e i cappelli vengono riconsegnati al buio. Qual'è la probabilità che nessuno ritorni in possesso del proprio cappello?).

PROB 24.[#] Nel preparare N panettoni, si immettono nell'impasto n acini di uvetta sultanina. Qual'è la probabilità che uno qualunque dei panettoni contenga almeno un acino?

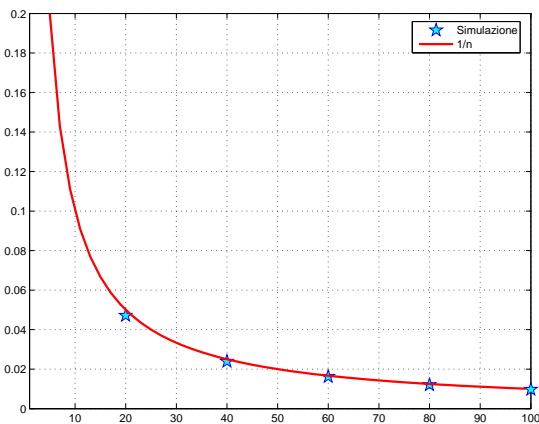
Soluzioni

#1 Le disposizioni dei tre dadi che portano a una somma uguale a 11 sono le 6 seguenti: $6+4+1$, $6+3+2$, $5+5+1$, $5+4+2$, $5+3+3$, $4+4+3$, ma tenendo conto delle permutazioni possibili ($6+4+1$, $6+1+4$, $4+1+6$, $4+6+1$, $1+6+4$, $1+4+6$ etc.) si tratta in realtà **27** su 216 possibilità, mentre per il 12 si ha $6+5+1$, $6+4+2$, $6+3+3$, $5+5+2$, $5+4+3$, $4+4+4$, sempre 6 partizioni, ma solo **25** combinazioni. DeMère si era ingannato non tenendo conto delle permutazioni possibili.

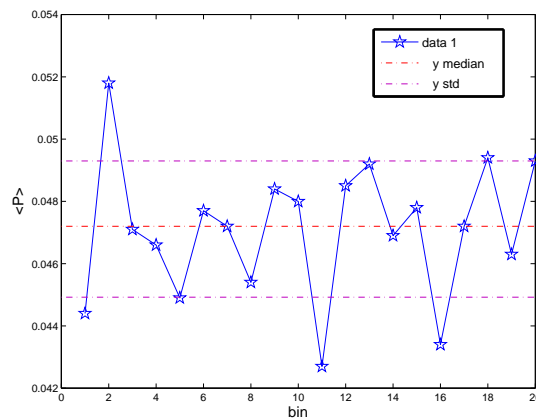
#3 Il problema si imposta facilmente in termini aritmetici. Se $n = c_0 + 10c_1 + 100c_2$ con c_i compreso tra 0 e 9, si avrà $n^3 = c_0^3 + 30c_0^2c_1 \pmod{100}$ (in quanto tutti gli altri termini comportano un fattore almeno pari a 100 e quindi non influiscono sulle ultime due cifre; 1000 non è compreso, ma è comunque da scartare). Ora il termine $30c_1^2c_2$ influenza solo la penultima cifra del risultato e quindi si dovrà avere $c_0^3 = 1 \pmod{10}$, che implica $c_0 = 1$ e di conseguenza $c_1 = 7$. Di tutte le scelte possibili per c_0 e c_1 solo una soddisfa la richiesta, dunque la probabilità è uguale a $1/100$. Un semplice comando matlab trae dall'impiccio nel caso non si sia particolarmente forti in aritmetica:

```
>> n=1:1000;
>> N3=n.^3;
>> length(find(N3 - floor(N3/100)*100 == 11))/1000
```

#4 Il problema rientra nella categoria generale delle “camminate aleatorie” (o random walks). In attesa di sviluppare qualche metodo di analisi, si può tentare una soluzione empirica basata sulla simulazione della coda. Il programma `rw.m`, in linguaggio matlab realizza una successione di situazioni a caso per un numero qualunque ($2n$) di spettatori.



(a) Stima di probabilità per il Probl. 4



(b) Media e fluttuazioni della simulazione della coda al botteghino

FIGURA 5-7. Dati dalla simulazione del problema della coda al botteghino

Esaminando l'output del programma per vari valori di n , si trova il diagramma di Fig. 7(a). Risulta naturale congetturare che la probabilità sia almeno approssimativamente $1/n$. La simulazione si realizza così: ogni spettatore con biglietto da 5 aumenta la scorta della cassa, mentre uno con biglietto da dieci la fa diminuire. Dunque una qualunque coda si può mettere in corrispondenza biunivoca con una successione di $+1$ e -1 in quantità uguali a n . Si ottengono tutte le possibili code considerando tutte le combinazioni di $2n$ elementi a n a n (che sono in numero di $\binom{2n}{n}$). Non è consigliabile tentare di enumerare tutte le code possibili e contare per quante di queste la cassa rimane sempre rifornita, per $n=100$ si tratta di un numero a 58 cifre! Però c'è il modo di estrarre a caso un numero ragionevole di combinazioni e sulla base di questo campionamento effettuare una stima. Naturalmente, dato che si tratta di una sorta di sondaggio, bisognerà tenere presente che il risultato è soggetto a errore statistico, e possibilmente bisognerà quantificarlo. La parte più difficile della simulazione è costituita di solito proprio dalla valutazione corretta delle fluttuazioni statistiche. Nella fig. 7(b) una simulazione per $n=20$ con un totale di 2×10^5 campioni (sempre poco numerosi rispetto al totale delle possibilità) è analizzata suddividendo il campione in gruppi di 20 e valutando la deviazione standard. Risulta che il valore più probabile della media è $1/21$, ossia in generale $1/(n+1)$. La soluzione si può determinare con il metodo della riflessione speculare (si veda [Gne79], Cap.I, Es. 5), oppure con il metodo della funzione generatrice, che ora utilizziamo. Sia $N(n)$ il numero di casi favorevoli (nessuno tra i $2n$ spettatori è costretto ad attendere). Sia poi $N_0(n)$ il numero di casi in cui la cassa rimane sempre in possesso di almeno un biglietto da 5 Euro, ad eccezione del momento iniziale e di quello finale, in cui per ipotesi non vi sono biglietti da 5 disponibili. Allora siano $2m$ gli spettatori che passano dalla cassa prima che questa sia priva di biglietti da 5; ovviamente $1 \leq m \leq n$ e si avrà

Dal grafico in Fig. 5-8 si capisce anche che ad ogni evento favorevole si può associare un cammino lungo due passi in più che ritorna a zero per la prima volta dopo $2(n+1)$ passi (il problema del “primo ritorno” di un cammino aleatorio) Abbiamo perciò che $N_0(n+1) = N(n)$ e quindi vale una relazione di ricorrenza che permette di ricostruire facilmente $N(n)$:

Sia $F(x) = \sum_{n \geq 0} N(n)x^n$; moltiplicando i due termini della relazione di ricorrenza per x^{n+1} si trova

Si ottiene così $F(x) = \frac{1-\sqrt{1-4x}}{2x}$. I coefficienti dello sviluppo in serie si possono ricondurre a $\frac{1}{n+1}\binom{2n}{n}$ e sono noti come numeri di Catalan [Sta99]. Il risultato dimostra che la probabilità richiesta dal problema è esattamente $1/(n+1)$.

- 218 -

#7 Se indichiamo con $p(x)$ la probabilità della “rovina” per un capitale iniziale x . Al primo tentativo, dopo una vincita la nuova probabilità sarà $p(x+1)$ e in caso di perdita $p(x-1)$. I due casi sono equiprobabili, perciò

$$p(x) = \frac{1}{2}(p(x+1) + p(x-1))$$

Dobbiamo allora cercare una soluzione a questa equazione con le condizioni $p(0) = 1$ e $p(m) = 0$. L'equazione è lineare e ammette due soluzioni indipendenti $p = 1$ e $p = x$. La soluzione va quindi cercata come $p(x) = \alpha + \beta x$. Imponendo le due condizioni, la soluzione è $p(x) = 1 - x/m$. Questo implica, ad esempio, che se si vuole raddoppiare il portafoglio tanto vale puntare tutto su una singola giocata, l'unica differenza essendo il tempo trascorso a giocare. Il programma `gambler.m` simula una giocata a caso e permette di valutare qualcosa in più rispetto alla analisi svolta fin qui, ad esempio risponde alla domanda “dato x e m quanto tempo in media passa prima della rovina?”. Risposta: $t = x(m - x)$.

#8 Dalla formula (5.1) si trova subito che $\mathcal{P}(\text{bianca}) = \frac{1}{10}(\frac{9}{2} + \frac{5}{6}) = \frac{32}{60}$. Dalla formula di Bayes si ottiene poi

$$\mathcal{P}(B_{10}|A) = \frac{\frac{1}{10} \times \frac{5}{6}}{\frac{32}{60}} = \frac{5}{32}.$$

#12 Si può determinare la funzione \mathcal{P} tenendo conto della relazione di ricorrenza

$$\mathcal{P}_{t+1}(n_1, n_2) = \frac{1}{4}(\mathcal{P}_t(n_1 + 1, n_2) + \mathcal{P}_t(n_1 - 1, n_2) + \mathcal{P}_t(n_1, n_2 + 1) + \mathcal{P}_t(n_1, n_2 - 1))$$

Si ha allora

$$\mathcal{P}_{t+1}(n_1, n_2) - \mathcal{P}_t(n_1, n_2) = \frac{1}{4}(\mathcal{P}_t(n_1 + 1, n_2) + \mathcal{P}_t(n_1 - 1, n_2) + \mathcal{P}_t(n_1, n_2 + 1) + \mathcal{P}_t(n_1, n_2 - 1) - 4\mathcal{P}_t(n_1, n_2))$$

e il membro di destra è l'analogo alle differenze finite dell'operatore di Laplace. Se introduciamo unità di lunghezza e di tempo rispetto a cui $t = T/\tau$ e $x = (n_1, n_2) = (X, Y)/a$ la formula precedente diventa

$$(\mathcal{P}_{T+\tau}(X, Y) - \mathcal{P}_T(X, Y))/\tau = \frac{a^2}{4\tau}(\mathcal{P}_T(X + a, Y) + \mathcal{P}_T(X - a, Y) + \mathcal{P}_T(X, Y + a) + \mathcal{P}_T(X, Y - a) - 4\mathcal{P}_T(X, Y))/a^2$$

Passiamo al limite per $a \rightarrow 0$ e $\tau \rightarrow 0$ in modo però che il rapporto $a^2/\tau = 2D$ si mantenga costante, allora l'equazione converge a

$$\frac{\partial \mathcal{P}_T(X, Y)}{\partial T} = \frac{D}{2} \Delta \mathcal{P}(X, Y)$$

dove Δ è l'operatore di Laplace. Il modello di random walk rappresenta perciò la discretizzazione del processo di diffusione. Determinare i valori medi $\mathcal{E}(X)$ e $\mathcal{E}(Y)$ in funzione del tempo.

#14 Sia $f_j(t)$ la probabilità che il processo torni per la prima volta in j dopo t passi. Allora si avrà

$$(5.1) \quad \begin{aligned} (P^t)_{jj} &= f_j(t) + f_j(t-1) P_{jj} + f_j(t-2) (P^2)_{jj} \\ &+ \dots + f_j(t-k) (P^k)_{jj} + \dots + f_j(1) (P^{t-1})_{jj} \\ &= \sum_{k=0}^{t-1} f_j(t-k) (P^k)_{jj} \end{aligned}$$

che si giustifica considerando che lo stato può tornare per la prima volta in k passi per poi tornare di nuovo (senza restrizioni) nei $t-k$ passi residui.

L'equazione permette di calcolare f_j dalla conoscenza delle potenze di P , un problema puramente algebrico. Dato che l'equazione è data da una convoluzione è facile risolverla introducendo le funzioni generatrici

$$(5.2) \quad \mathfrak{G}_j(s) = \sum_{t=1}^{\infty} (P^t)_{jj} s^t$$

$$(5.3) \quad \mathcal{F}_j(s) = \sum_{t=1}^{\infty} f_j(t) s^t.$$

Dall'Eq.(5.1) si trova allora

$$\mathfrak{G}_j(s) = \mathcal{F}_j(s) + \mathcal{F}_j(s) \mathfrak{G}_j(s)$$

ossia

$$(5.4) \quad \mathcal{F}_j(s) = \frac{\mathfrak{G}_j(s)}{1 + \mathfrak{G}_j(s)}.$$

Dalla soluzione possiamo verificare che $\mathcal{F}_j(1)$, ossia la probabilità totale di tornare in j , risulta i) minore di uno se la serie $\mathfrak{G}_j(1)$ è convergente, ii) uguale a uno se la stessa serie diverge. Inoltre il tempo medio di ritorno, ossia $\tau_j = \sum_{t=1}^{\infty} t f_j(t)$ risulta dato da $\tau_j = \frac{d\mathcal{F}_j}{ds}|_{s=1}$. Sia u_j la distribuzione stazionaria, soluzione di $Pu = u$. Allora si dimostra che vale il risultato semplicissimo (vedi Appendice 5.8)

$$(5.5) \quad \tau_j = 1/u_j.$$

Nel caso del modello di Ehrenfest, per il quale $\mu_j = \binom{N}{j}/2^N$, il tempo medio di ritorno è allora 2^N per lo stato in cui tutte le particelle sono inizialmente collocate in una sola urna.

#15 Lo spettro della matrice

$$A = \begin{pmatrix} -\alpha & 0 & \gamma \\ \alpha & -\beta & 0 \\ 0 & \beta & -\gamma \end{pmatrix}$$

è dato dalle radici dell'equazione $(\lambda + \alpha)(\lambda + \beta)(\lambda + \gamma) = \alpha\beta\gamma$, il che implica che una radice è zero e le rimanenti due hanno parte reale negativa. Il processo converge dunque alla distribuzione asintotica

$$p_A = \mathcal{N}\beta\gamma, p_B = \mathcal{N}\gamma\alpha, p_C = \mathcal{N}\alpha\beta$$

che rappresenta l'autovettore $Ap = 0$. \mathcal{N} è fissato dalla normalizzazione a uno della probabilità totale.

#19 Si vedano [Wig67, Meh89]. Il problema si può esplorare “sperimentalmente” con *matlab*:

Wigner-caso simmetrico reale

```
function E=wigner(N, samples)
E=[];
for n=1:samples
    M = randn(N); % distrib. normale per gli elementi di matrice
    M=(M+M')/2;   % simmetrizzazione
    E=[E; eig(M)]; % accumula gli autovalori
end
hist(E, samples*N/1000); % disegna l'istogramma
```

Si noterà che gli autovalori si distribuiscono secondo una densità concentrata nell'intervallo $[-\sqrt{2N}, \sqrt{2N}]$. Ciò si ripete nel caso di matrici hermitiane

Wigner-caso Hermitiano

```
function E=wigner(N, samples)
E=[];
for n=1:samples
M = randn(N)+i*randn(N);
M=(M+M')/2/sqrt(2);
E=[E; eig(M)];
end
hist(E, samples*N/1000);
```

Un piccolo numero di esperimenti può essere sufficiente per stabilire la legge generale. Per la teoria si vedano i riferimenti consigliati.

#20 Il caso unitario è incluso negli studi di Mehta et al, quindi i riferimenti sono gli stessi. Per eseguire esperimenti numerici si può utilizzare un algoritmo non banale[†]

Unitary Ensemble (ue.m)

```
function U=ue(N)
% Return a random unitary matrix NxN
% uniformly distributed according to Haar measure on U(N)
% Algorithm by F. Mezzadri, Bristol U. 2006
Z = randn(N)+i*randn(N);      % Gaussian complex
[Q,R] = qr(Z);                % QR-decomposition
D = diag(R);
D = diag(D./abs(D));          % (1,-1) diagonal
U = Q*D;                      % return the unitary sample
```

#21 L'algoritmo è implementato in *matlab* come segue (Onofri-Destri, unpublished). Il programma mostra come lo spettro, oltre all'autovalore uno comune a tutte le matrici stocastiche, riempie il disco di raggio $1/\sqrt{N}$; la distribuzione è uniforme nel disco tranne che per un addensamento sul segmento $[-1,1]$. Se trovate una dimostrazione formale, pubblicatela.

rsm.m

```
function E=rsm(N, sample)
% Usage:
% E=rsm(N, sample, col)
% N=dimension, sample=number of random matrices
phi=linspace(0,2*pi,100);
plot(exp(i*phi), 'r', 'LineWidth', 2);
hold on
xi = -log(rand(N));           % log-uniform
x=xi./(ones(max(size(xi)),1)*sum(xi)); % normalizzare
E=eig(x);
h=plot(E*sqrt(N), '.', 'MarkerSize', 1, 'EraseMode', 'none');
axis([-1.1,1.1,-1.1,1.1]);
axis square
for j=1:sample
xi=-log(rand(N));
x=xi./(ones(max(size(xi)),1)*sum(xi));
W=eig(x);
set(h, 'XData', real(W)*sqrt(N), 'YData', imag(W)*sqrt(N))
drawnow
E=[E; W];
end
```

[†]Ringrazio vivamente F. Mezzadri per avermi fornito il codice.

#23 Per fissare le idee: S_3 è costituito dalle sei permutazioni

$$\begin{pmatrix} 123 \\ 123 \end{pmatrix} \begin{pmatrix} 123 \\ 312 \end{pmatrix} \begin{pmatrix} 123 \\ 231 \end{pmatrix} \begin{pmatrix} 123 \\ 132 \end{pmatrix} \begin{pmatrix} 123 \\ 213 \end{pmatrix} \begin{pmatrix} 123 \\ 321 \end{pmatrix}$$

dunque $D(3) = 1/3$. Ma il numero di permutazioni è $n!$ e non è quindi il caso di procedere nello stesso modo per $n = 16$ che comporta circa 1.3×10^{12} permutazioni. Utilizzando la funzione

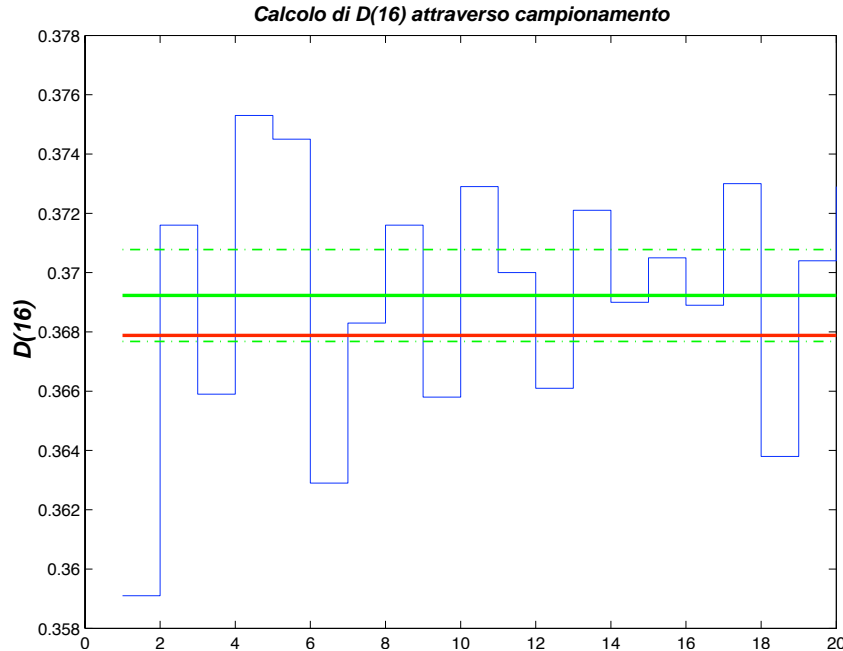


FIGURA 5-9. Stima del numero di permutazioni senza punto fisso.

`randperm` già nota dall'Eserc. 5-1 a Pag. 183, è molto facile farsi un'idea di quanto può valere $D(16)$. Secondo la simulazione (20 pacchetti di 10000 permutazioni scelte a caso, molto inferiore al numero totale di permutazioni possibili) si ha che $D(16) \approx .369 \pm 0.004$ (Fig. 5-9). Il codice sta, volendo, in una riga:

```
function p = derangements(N, smpl)
n=0; for j=1:smpl, if(any(randperm(N)==1:N)), n=n+1; end; end;p=n/N;
```

Vediamo cosa ci può dire la formula di inclusione/esclusione. Partiamo dal fatto che per ogni i ($1 \leq i \leq n$) l'insieme delle permutazioni che lascia fisso i è un gruppo di permutazioni su $n-1$ elementi, indichiamolo con A_i . Inoltre $A_i \cap A_j$ è l'insieme di permutazioni che lascia due elementi fissi, e così via. Avremo perciò (indicando con $|A|$ il numero di elementi di A)

$$\begin{aligned} n! D(n) &= n! - \sum_i |A_i| + \sum_{i>j} |A_i \cap A_j| - \sum_{i>j>k} |A_i \cap A_j \cap A_k| + \dots \\ &= n! - \binom{n}{1} (n-1)! + \binom{n}{2} (n-2)! - \dots + (-1)^k \binom{n}{k} (n-k)! + \dots + (-1)^n \\ &= \sum_{k=0}^n (-1)^k \frac{n!}{k!(n-k)!} (n-k)! \end{aligned}$$

per cui

$$D(n) = \sum_{k=0}^n (-1)^k \frac{1}{k!}.$$

Per n grande si ha perciò $D(n) \approx \exp(-1) \approx .3679$. Il problema era stato risolto già da Poincaré [Poi12] dove si trova anche la formula di inclusione/esclusione (vedi Pag. 186).

#24 Considerando i chicchi di volume trascurabile rispetto a quello dei panettoni, possiamo considerare che la distribuzione delle uvette sia uniforme. La presenza di una uvetta in un dato panettone sarà $1/N$. Il problema è quindi equivalente a una successione di n estrazioni di Bernoulli con $p = 1/N$, e quindi $\mathcal{E} = np = n/N$. Dato che p è presumibilmente piccolo e sia n che N sono grandi, la distribuzione binomiale converge a quella di Poisson e la probabilità cercata è data da $1 - \mathcal{P}[0] = 1 - \exp\{-np\} = 1 - \exp\{-n/N\}$.

Bibliografia

- [AO84] G.E. Andrews and E. Onofri, *Lattice gauge theory, orthogonal polynomials, and q -hypergeometric functions*, Special Functions: Group Theoretical Aspects and Applications (A. Askey, T. H. Koornwinder, and W. Schempp, eds.), D. Reidel Pub.Co., 1984, pp. 163–188.
- [Arn78] V. Arnol'd, *Metodi matematici della meccanica classica*, Editori Riuniti, Roma, 1978.
- [Bar47] V. Bargmann, *Irreducible Unitary Representations of the Lorentz Group*, Ann. Math, **48** (1947), no. 3, 568.
- [BCO07] M. Bonini, G. M. Cicuta, and E. Onofri, *Fock space methods and large N* , J. Phys. **A40** (2007), F229–F234.
- [BCR84] K. Baclawski, M. Cerasoli, and G.C. Rota, *Introduzione alla Probabilità*, UMI, 1984.
- [Bel87] J. S. Bell, *Speakable and unspeakable in quantum mechanics*, Cambridge UP, Cambridge, 1987.
- [BSRW02] M. Bennet, M/F/ Schatz, H. Rockwood, and K. Wiesenfeld, *Huygen's clocks*, Proc. R. Soc. Lond. **A458** (2002), 563–579.
- [BW69] C. M. Bender and T. T. Wu, *Anharmonic Oscillator*, Phys. Rev. **184** (1969), no. 5, 1231–1260.
- [CH62] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Interscience, 1962.
- [CMO80] M. Casartelli, G. Marchesini, and E. Onofri, *A singular integral operator arising from $1/N$ expansions: analytical and numerical results*, J.Phys.A: Math.& Gen. **13** (1980), 1217–1225.
- [CP78] F. Calogero and A.M. Perelomov, *Asymptotic density of the zeros of hermite polynomials of diverging order, and related properties of certain singular integral operators*, Lett.Nuovo Cim. **23** (1978), 650.
- [DH87] P.H. Damgaard and H.Hüffel, *Stochastic Quantization*, Phys.Rep. (1987), no. 152(5-6), 227.
- [Dir29] P.A.M. Dirac, *The Principles of Quantum Mechanics*, Clarendon Press, Oxford, 1929, traduzione italiana ediz. Boringhieri, Torino, 1959.
- [EHW75] J. Ehlers, K. Hepp, and H.A. Weidenmüller (eds.), *Dynamical Systems, Theory and Applications*, Lecture Notes in Physics, vol. 38, Berlin, Springer, 1975.
- [Erd55] A. Erdélyi (ed.), *Higher Transcendental Functions*, vol. I-III (Bateman Manuscript Project), McGraw-Hill, New York, 1955.
- [FDO04] V.A. Fateev, R. DePietri, and E. Onofri, *Exact and semiclassical approach etc.*, J.Phys.A **37** (2004), 11379–11390.
- [Fel68] W. Feller, *An Introduction to Probability Theory and its Applications*, Wiley series in probability and mathematical statistics, J. Wiley & Sons, New York, 1968.
- [Fey72] R.P. Feynman, *Statistical Mechanics*, W. A. Benjamin, Inc., Reading, Mass., 1972.
- [FFS82] M.D. Feit, J.A. Fleck, and A. Steiger, *Solution of the Schroedinger equation by a spectral method*, J. Comput. Phys. **47** (1982), 412–433.
- [FH65] R.P. Feynman and A.R. Hibbs, *Quantum mechanics and path integrals*, McGraw-Hill, 1965.
- [Fis63] M. Fisz, *Probability Theory and Mathematical Statistics*, 3rd ed., J. Wiley & Sons, New York, 1963.
- [Gan59] F.R. Gantmacher, *The Theory of Matrices*, Chelsea Pub. Co., New York, 1959.
- [Gel89] I. M. Gelfand, *Lectures on linear algebra*, Dover, 1989.
- [Gin00] M. Gintis, *Game Theory Evolving*, Princeton UP, 2000.
- [GL74] I. Glazman and Y. Liubitch, *Analise Linéaire dan les Espaces de Dimension Finies*, MIR, Moscou, 1974.
- [GL96] G.H. Golub and C.F. Van Loan, *Matrix Computations*, iii ed., Johns Hopkins U.P., Baltimore, USA, 1996.
- [GMS94] M. Goossens, F. Mittelbach, and A. Samarin, *The L^AT_EX Companion*, Addison-Wesley, Reading, Mass., 1994.

- [Gne79] B. Gnedenko, *Teoria della probabilita'*, MIR, 1979.
- [GR65] L.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products*, IV ed., Academic Press, New York, 1965.
- [GW69] G.H. Golub and J.H. Welsch, *Calculation of gauss quadrature rules*, Math. Comput. **23** (1969), 221–230.
- [Hal82] P.R. Halmos, *A Hilbert space problem book*, Springer, N.Y., 1982.
- [Ham62] M. Hamermesh, *Group Theory and its Applications to Physical Problems*, Addison Wesley, Reading, Mass., 1962.
- [Hel78] S. Helgason, *Lie Groups and Symmetric Spaces*, Academic Press, New York, 1978.
- [HLP52] G. Hardy, J.E. Littlewood, and G. Polya, *Inequalities*, Cambridge U.P., 1952.
- [Hoc71] H. Hochstadt, *The Functions of Mathematical Physics*, Wiley-Interscience, New York, 1971.
- [ID89] C. Itzykson and J. M. Drouffe, *Théorie Statistique des Champs*, InterEditions, Paris, 1989.
- [Kac59] M. Kac, *Probability and Related Problems in the Physical Sciences*, Interscience, New York, 1959.
- [Kir74] A. Kirillov, *Elements de la Théorie des Représentations*, MIR, Mosca, 1974.
- [Knu98] D. Knuth, *The art of computer programming*, Addison–Wesley, 1998.
- [KP99] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Applications of Mathematics, vol. 23, Springer, Berlin, 1999.
- [KS85] J.R. Klauder and B. Skagerstam (eds.), *Coherent States, Applications in Physics and Mathematical Physics*, World Scientific, Singapore, 1985.
- [Lan50] C. Lanczos, *An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Equations*, J. Res. Nat. Bureau St. **45** (1950), no. 4, 255–282.
- [Lit50] D.E. Littlewood, *The Theory of Group Characters*, Oxford U.P., 1950.
- [LL76] L.D. Landau and E.M. Lifshitz, *Meccanica quantistica, teoria non relativistica*, Editori Riuniti, Roma, 1976.
- [Mad70] I. J. Maddox, *Elements of Functional Analysis*, Cambridge University Press, Cambridge, 1970.
- [Mal03] G. Maltese, *Enrico Fermi in America*, Le Ellissi, Zanichelli, Bologna, 2003.
- [Meh89] M.L. Mehta, *Matrix Theory, Selected Topics and Useful Results*, Les Editions de Physique, Les Ulis (F), 1989.
- [ML78] C. Moler and C. Van Loan, *Nineteen Dubious Ways To Compute The Exponential Of A Matrix*, SIAM Review **20** (1978), no. 4, 801–836.
- [Mos65] F. Mosteller, *Fifty challenging problems in probability with solutions*, Dover, New York, 1965.
- [Nai68] M. A. Naimark, *Linear Differential Operators*, vol. 2, Frederic Ungar, New York, 1968.
- [Nel67] E. Nelson, *Dynamical Theories of Brownian Motion*, Princeton University Press, Princeton, N.J., 1967.
- [New66] R.G. Newton, *Scattering Theory of Waves and Particles*, McGraw-Hill, New York, 1966.
- [NO76] A. Nikiforov and V. Ouvarov, *éléments de la Théorie des Fonctions Speciales*, MIR, Mosca, 1976.
- [OD96] E. Onofri and C. Destri, *Istituzioni di fisica teorica*, La Nuova Italia Scientifica, 1996.
- [OD99] E. Onofri and C. Destri, *Istituzioni di Fisica Teorica*, Carocci, Roma, 1999.
- [OMT91] E. Onofri, P. Maraner, and G.P. Tecchiolli, *Spectral Methods in Computational Quantum Mechanics*, Int. J. Comput. Appl. Math. (1991), 209–219.
- [Ono84] E. Onofri, *Lezioni sulla teoria degli operatori lineari*, Ediz. Zara, Parma, 1984.
- [Ono85] ———, *A nonlinear variational problem*, SIAM Rev, 1985, 27.
- [Pel94] A. Pelissetto, *Elementary particles, quantum fields and statistical mechanics*, Elementary particles, Quantum Fields and Statistical Mechanics (M. Bonini, G. Marchesini, and E. Onofri, eds.), Università di Parma, 1994.
- [Pen89] R. Penrose, *The Emperor New Mind*, Oxford U. Press, 1989.
- [Pen97] ———, *La mente nuova dell'imperatore*, BUR, Rizzoli, Milano, 1997.
- [Poi12] H. Poincaré, *Calcul des Probabilités*, Gauthiers–Villars ed., (disponibile su *gallica.fr*), Paris, 1912.
- [Pol77] A. M. Polyakov, Nucl.Phys. **120B** (1977), 429.
- [PR69] Yu.V. Prohorov and Yu.A. Rozanov, *Probability Theory*, Springer Verlag, Berlin, 1969.

- [Put67] C.R. Putnam, *Commutation Properties of Hilbert Space Operators and related topics*, Erg.d.Math.und i.Grenz., vol. 36, Springer-Verlag, New York, 1967.
- [PW81] G. Parisi and Y.-S. Wu, *perturbation theory without gauge fixing*, Sci. Sinica (1981), 483.
- [QSS00] A. Quarteroni, R. Sacco, and F. Saleri, *Matematica Numerica*, Springer-Verlag Italia, Milano, 2000.
- [QSS07] ———, *Numerical Mathematics*, 2nd ed., Texts in Applied Mathematics, vol. 37, Springer, 2007.
- [Reb83] C. Rebbi, *Lattice Gauge Theory and Montecarlo simulations*, World Sci., Singapore, 1983.
- [Roz69] Yu. A. Rozanov, *Probability theory: a concise course*, Dover, 1969.
- [RS78] M. Reed and B. Simon, *Methods of Mathematical Physics*, vol. II, Academic Press, New York, 1978.
- [RS80] ———, *Methods of Mathematical Physics: I - Functional Analysis*, vol. I, Academic Press, New York, 1980.
- [Sak90] J.J. Sakurai, *Meccanica quantistica moderna*, Zanichelli, Bologna, 1990.
- [Shi77] G. Shilov, *Linear algebra*, Dover, 1977.
- [Sok91] A. D. Sokal, *How to beat critical slowing down: 1990 update*, Lattice 90 (U.M.Heller et al, ed.), North-Holland, 1991.
- [Sok97] ———, *Functional integration (Cargèse, 1996)*, ch. Monte Carlo methods in Statistical Mechanics: Foundations and New Algorithms, pp. 131–192, Plenum, New York, 1997.
- [Sta99] R. P. Stanley, *Enumerative Combinatorics*, vol. 2., Cambridge U.P., 1999.
- [Ste94] S. Sternberg, *Group Theory and Physics*, Cambridge U.P., 1994.
- [Sze59] G. Szegő, *Orthogonal Polynomials*, American Mathematical Society, Providence, R.I., 1959.
- [Tay58] A.E. Taylor, *Introduction to functional Analysis*, J. Wiley, New York, 1958.
- [TN02] E.O. Tuck and J.N. Newman, *Longitudinal waves in slender moonpools*, 17th Int. Workshop on Water Waves and Floating Bodies, Cambridge (London) (R. Rainey, ed.), RINA, 2002.
- [Tri85] F.G. Tricomi, *Integral Equations*, Dover Pub.s, New York, 1985.
- [Var62] R. S. Varga, *Matrix iterative analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.
- [VW06] G. Veneziano and J. Wosiek, *Planar quantum mechanics: An intriguing supersymmetric example*, JHEP **01** (2006), 156.
- [Wax54] N. Wax (ed.), *Selected papers on Noise and Stochastic Processes*, Dover, New York, 1954.
- [Wic55] D. Wick, *The infamous boundary: seven decades of heresy in quantum physics*, Birkhäuser, Boston, 1955.
- [Wig67] E. P. Wigner, *Random Matrices in Physics*, SIAM Review **9** (1967), no. 1, 1–23.

Indice analitico

Symbols

$C_2(a, b)$	111
H_2	107
$L_2(a, b)$	111
$\ \cdot\ $	108
\mathbb{I}	36
$\mathbf{C}\sigma$	124
$\mathbf{P}\sigma$	124
$\ker()$	96, 113
p	128
q	128
ad_X	58, 103
\mathfrak{B}	151
$\mathfrak{B}(\mathcal{C})$	37
$\langle \mid$	32
\mathbb{C}^n	27
$\mathbb{D}[\cdot]$	viii
Δ	147
E_λ	48
ε_{ijk}	31 <i>n</i>
$\hat{f}(\mathcal{X})$	43
$\mid \rangle$	32
ℓ_2	107, 109
$M(n, \mathbb{R})$	24
$M[\cdot]$	viii
$[n]!$	85
$\ \cdot\ $	38
\mathbb{P}_λ	48
$R(n, \alpha)$	35
$\text{RES}()$	124
$SU(2)$	30
$S[\cdot]$	187
σ_i	30
$\text{Tr}()$	84
$n \neq \infty$	ix

A

a.c.	<i>vedi</i> funzione, assolutamente continua
<code>absp.m</code>	159
aggiunto Hermitiano	52 , 117
aleatoria	<i>vedi</i> variabili, aleatorie

algebra	
degli operatori	37
alias	
metodo degli	198
Andrews	162, 168
aquamacs-emacs	ii
Armoniche sferiche	100
Arnold	79
arpack	144
autoaggiunto	<i>vedi</i> operatore, autoaggiunto
autocorrelazione	199

B

b.c.(boundary conditions)	<i>vedi</i> c.c.
Bargmann	168 , <i>vedi</i> spazio, di Bargmann
rappresentazioni di $SU(1,1)$	162
base	24
Bayes	<i>vedi</i> formula, di Bayes
BCH	<i>vedi</i> formula, di
Baker–Campbell–Hausdorff	
Bell	182
Bender-Wu	60
Bergman	<i>vedi</i> nucleo di Bergman
Bernoulli	168 , 188, 190, <i>vedi</i> distribuzione, di Bernoulli
Bessel	<i>vedi</i> disuguaglianza, di Bessel
bessel.m	165
binomiale	<i>vedi</i> distribuzione, binomiale
Bonini	162
bra	<i>vedi</i> Dirac, notazione di

C

c.c.	10, 116, 120–124, 150, 154, 156
cambiamenti di base	36
canonici	<i>vedi</i> operatori, canonici
Cartan	168
Casartelli	98, 163
catene di Markov	195
Cauchy	169 , <i>vedi</i> successioni, di Cauchy
integrale di	43
Cayley	

matrici di	30
teorema di	49
trasformata di	136
chiuso	<i>vedi</i> operatori, chiusi
chiusura	
di un operatore	115
Christoffel–Darboux	
formula di	89
Cicuta	162
circolante	
matrice	88
code	213
codici . 7, 11, 62, 96–98, 103, 126, 144, 147, 159,	
163, 183, 194, 197, 205, 208, 217, 220–222	
codominio	114
compatto	
operatore	135
completamento	110
completezza	111
completo	
spazio lineare	108
condizionata	<i>vedi</i> probabilità condizionata
contorno	
condizioni al	<i>vedi</i> c.c.
corda vibrante	13
correlazione	189
Courant	136
Criterio di Schur	116

D

decomposizione spettrale	45, 192
$\delta(f)$	151, 152
$\delta(x)$	113, 133
DeMère	213
derangements.m	222
determinante	
di Gram	76
di Vandermonde	81, 95
DFT	84
diagonalizzabilità	
teorema di	49
difetto	
indici di	122, 136
differenziali	<i>vedi</i> operatori, differenziali
dimensionalità	24
Dirac	169
funzione $\delta(x)$	113, 133
notazione di	32–34
Dirichlet	169
condizioni di	123
distribuzione	
binomiale	184, 190, 203
di Gauss	203
di Bernoulli	184, 203

di Gauss multidimensionale	203
di Poisson	203
di Polya	205
di probabilità	187
normale	203
distribuzione stazionaria	192
disuguaglianza	
di Holder	77
di Minkowski	77
triangolare	77
di Bessel	108
di Cauchy–Schwartz	77
di Young	79, 133, 134
disuguaglianze	76–79
doppio prodotto vettoriale	17
Drouffe	193
dueosc.m	7
Dyson	169

E

Ehrenfest	<i>vedi</i> modello, di Ehrenfest
eigs	144
emacs	ii
ensemble unitario	221
equazione	
di Kolmogorov	209
di Langevin	200
Erdelyi	75, 76
Esempi . . 36, 43, 57, 61, 63, 65, 70, 74, 108, 111,	
113, 114, 117, 118, 125, 126, 128, 132, 137,	
146, 184, 188, 189	
estensione	
di operatore	115
operatori simmetrici	136
Euler	169

F

Famiglia spettrale	130
Fateev	163
Feller	182
Fermi	170, 196
Feynman	141, 142, 170, 209
FFTW	144n
fftn	147
Fisz	206
Flaschka	88
formula	
di Baker–Campbell–Hausdorff	58
di Bayes	186
di Lie	57
di Riesz–Dunford	43
di Lie–Trotter	141–144
Fourier	170

trasformata	13, 139
discreta	84
veloce	143
Frobenius ... <i>vedi</i> teorema, di Perron–Frobenius	
funzionale	
limitato	112
funzionali lineari	112
funzione	
assolutamente continua	121
di Green	14–15, 150, 159
generatrice	220
Funzione di distribuzione	188
funzioni di operatori	37–41
Funzioni di più operatori	55

G

Gambler's ruin	213
Gantmacher	23 <i>n</i> , 96, 204
Gauss	171 , <i>vedi</i> distribuzione, di Gauss
formule di quadratura	73–75
GaussHerm.m	75
GaussLeg.m	75
GaussHerm.m	75
Gelfand	23 <i>n</i>
Gershgorin	<i>vedi</i> teorema, di Gershgorin
teorema di	192
Gintis	214
giocatore	
problema del	213
Glazman	81
Golub	41, 52, 72, 96, 99, 144, 146, 171
Gradshteyn-Ryzhik	76, 98
grafico	
di un operatore	115
Green	<i>vedi</i> funzione, di Green

H

Halmos	149, 158
Hamermesh	101
Hamilton	30 <i>n</i> , 172
Hardy	79, 172
Heisenberg	170
regole di commutazione	128
Helgason	110, 111
Hermite	172 , <i>vedi</i> polinomi, di Hermite
Hilbert	172
spazio di	110
spazio di	107
Hochstadt	73, 75, 99
Holder	77
hpm.m	97
huygens.m	19

I

inclusione/esclusione	185
indice di autovalore	49
indipendenza	
statistica	187
indipendenza lineare	24
Insiemi completi di operatori	55
Insiemi irriducibili di operatori	56
integrale	
di Riesz-Dunford	43
inverso	
calcolo dell'	41
iperbolica	<i>vedi</i> matrice, iperbolica
Ising	<i>vedi</i> modello, di Ising
isometrico	<i>vedi</i> operatore, isometrico
Itzykson	182, 193

K

Kac	96, 173 , 182
ket	<i>vedi</i> Dirac, notazione di
Kirillov	90
Klauder	110, 166, 173
Kloeden-Platen	211
Knuth	ii, ix, 173 , 194
Kolmogorov 174 , <i>vedi</i> equazione, di Kolmogorov	

L

Lagrange	174
Laguerre	<i>vedi</i> polinomi, di Laguerre
Lamport	ii, 174
Lanczos	
algoritmo di	145
algoritmo di	99
Landau-Lifshitz	37
Langevin	<i>vedi</i> equazione, di Langevin
Legendre	174 , <i>vedi</i> polinomi, di Legendre
Lemma di Schur	56
Lie	174 , <i>vedi</i> formula, di Lie
Lie–Trotter	<i>vedi</i> formula, di Lie–Trotter
Littlewood	96

M

Maddox	112
Maltese	196
Markov	175 , <i>vedi</i> processi di Markov
catene di	191, 195
matrice	
circolante	88
iperbolica	ii, 87, 92, 98
irriducibile	214
metrica	27, 28, 69
rappresentativa	36

stocastica 100, 192
 unitaria 29
 matrice di trasformazione 25
 matrici
 random
 hermitiane 215
 positive 215
 unitarie 215
 sparse 144–147
 tridiagonali 89, 144
 Mehta 96, 98, **175**, 220
 metodo di “splitting” 141
 Metropolis 196
 Mezzadri 221 *n*
 modello
 di Ehrenfest 191
 di Ehrenfest
 soluzione 206
 continuo 215
 di Ising 199
 modi normali 6
 Moler 66, **175**
 MonteCarlo 196–201
 dinamico 198
 Moser 88

N

Naimark 131, 133, 135
 Nelson 141
 Neumann
 condizioni di 123
 Newton 118
 norma
 di un operatore 38
 normale *vedi* operatore, normale
 Notazioni viii
 nucleo di Bergman 110

O

operatore
 di proiezione 48
 autoaggiunto 52, 117
 chiuso **114**, 115
 compatto **135**
 continuo **114**
 di proiezione 35
 di shift 125
 illimitato 118
 inverso 41
 isometrico 117
 limitato **114**
 normale 52, **55**
 risolvente **42**, 44, 60, 124, 126, 133, 150

 proprietà 127
 unitario 52, **53**, **117**
 operatori
 canonici 128
 di rotazione 37
 differenziali 120, 133
 nello spazio di Hilbert 114–124
 oscillatori accoppiati 6

P

parametro di rilassamento 91
 Parisi 200
 Pascal **175**, 213
 Path-integrals 141
 Pauli **170**
 matrici di 30
 Pelissetto 198
 Penrose 182
 permutazioni 216
 Perron–Frobenius
 teorema di 192
 perturbazioni
 teoria delle 59
 peso 69
 Poincaré **176**, 186, 223
 Poisson *vedi* distribuzione, di Poisson
 poker 183 *n*
 polinomi
 classici 75–76
 di Hermite 76
 di Laguerre 76
 di Legendre 70, 76
 di Tchebychev 164
 ortogonali 69–76
 polinomio
 caratteristico 42
 minimale 49
 Polya **176**, 193, *vedi* distribuzione, di Polya
 Polyakov 209
 portalettere
 il problema del 197
 primo ritorno 218
 probabilità
 misura 183
 spazio di 182
 probabilità condizionata 186
 problemi 17–20, 30, 39, 41, 43, 48, 54, 72,
 81–103, 147–166, 183–197, 213–223
 processi di Markov 191
 processi diffusivi 209
 prodotto interno 27
 proiettore *vedi* operatore, di proiezione
 Putnam 128

Q

quadratura .. *vedi* Gauss, formule di quadratura
 quantizzazione stocastica 201
 Quarteroni 74, 192
 quaternioni 30, 30*n*

R

rsm.m 215, 221
 random variables 187
 random walk 100, 192, 213, 214
 range *vedi* codominio
 Rappresentazione di interazione 67
 Reed-Simon 134–138
 Reitter ii
 relazione
 di completezza 111
 relazione di ricorrenza 72
Resolvent.nb 126
 reticolo duale 147
 Ricci
 simbolo di 31*n*
 ricorrenza, relazioni di 71
 Riesz *vedi* teorema, di Riesz
 Riesz–Dunford .. *vedi* formula, di Riesz–Dunford
 risolvente *vedi* operatore, risolvente
 ritorno
 tempo medio di 207, 214
 ritorno all’origine 213
 Rodriguez
 formula di 75
 rotazioni 35

S

S2kit 144*n*
schr3D.m 146
 Schroedinger **176**
 Schur *vedi* Criterio di Schur
 lemma di 56
 separabile *vedi* spazio, separabile
 serie
 asintotica 59
 perturbativa 59
 Shilov 23*n*
 similitudine
 trasformazioni di 58
 Simon **177**
 sistemi lineari 66
 skewness 187
 somma diretta 111
 sottospazio 25
 spazi lineari 23
 spazi lineari complessi 27

spazio
 di Bargmann 151, 152, 160
 di Hilbert 107, 110
 di probabilità *vedi* probabilità, spazio di
 separabile 110
 spazio unitario 27
 spettro 44
 continuo 124
 discreto 124
 puntuale 124
 residuo 124
 Stallman ii, **177**
 Stanley 185, 218
 Sternberg 101
 Stone *vedi* teorema, di Stone
 Sturm
 criterio di 146*n*
 successioni
 pseudo casuali 196
 di Cauchy 108
susyqm.m 163
 SVD 175
 sviluppo di Dyson 67
 sviluppo di Lagrange 41
 sviluppo spettrale 48
 Szegö 75

T

Taylor 112, 116, 136
 tensore
 d’inerzia 3
 teorema
 di Wintner 128
 di Cayley 49
 di Gershgorin 204
 di Perron–Frobenius 204
 di Riesz 112
 di Stone 138
 teoria spettrale 42, 124–127
 Toda
 modello di 88*n*
 traccia
 di una matrice 84
 trasformata
 di Fourier 139
 trasformata di Cayley 136
 trasformata di Fourier
 discreta 84
 veloce 143
 trasformazioni di base 25
 Tricomi 158, 164
tridiag.c 146*n*
 tridiagonale *vedi* matrici
 Tuck **177**

operatore di 153*n*

U

ue.m 221

unitario *vedi* operatore, unitario
spazio 27

V

Varga 204

variabili

aleatorie 187

indipendenti 187

Veneziano 162

Volterra 67, 150, **178**

Von Neumann **177**

W

Wax 182

Wigner **178**, 220

legge del semicerchio 215

wigner.m 220

Wintner

teorema di 128

Y

Young *vedi* disuguaglianza, di Young

Z

Zeri

dei polinomi ortogonali 72

delle funzioni di Bessel 165

Elenco delle figure

1-1 Oscillatori accoppiati	6
1-2 n oscillatori accoppiati.	8
1-3 La corda vibrante.	13
1-4 Deformazione della fune sotto una forza applicata in un punto p .	14
2-1 Brougham Bridge, Dublin, dove i quaternioni sono venuti alla luce.	30
3-1 Calcolo di \expm con Mathematica (1).	40
3-2 Calcolo di \expm con Mathematica (2).	40
3-3 $\oint_{\partial\mathcal{U}} = \sum_k \oint_{\mathcal{H}_k}$	46
3-4 Dimostrazione grafica della disuguaglianza di Holder.	78
3-1 Il dado del Probl. 63	90
3-2 La trasformazione 3.1	101
4-1 La successione di Cauchy non convergente in C_2	108
4-2 Il cammino di integrazione Γ viene deformato in modo da adattarsi allo spettro sull'asse reale.	129
4-3 Il cammino di integrazione Γ nel calcolo di $f(p^2)$.	134
4-4 La formula di Lie–Trotter al lavoro.	143
5-1 Una partita a poker simulata.	185
5-2 Estratto da Poincaré, <i>Calcul des Probabilités</i>	186
5-3 De Moivre: legge dei grandi numeri, istogramma con $\operatorname{erfc}()$ sovrapposta	191
5-4 Distribuzione di Polya	206
5-5 Lo spostamento $x(t) \rightarrow x(t + dt)$ decomposto in parte casuale e in trascinamento (<i>drift</i>)	209
5-6 Il grafo del Probl. 13	214
5-7 Dati dalla simulazione del problema della coda al botteghino	217
5-8 Confronto tra cammini sempre positivi e cammini non–negativi	218
5-9 Stima del numero di permutazioni senza punto fisso.	222